PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/54, 9/12, C12Q 1/48
C12N 1/19 // C12Q 1/02
(C12N 1/19, C12R 1:645)

A1

(11) International Publication Number: WO 94/03609

(43) International Publication Date: 17 February 1994 (17.02.94)

(21) International Application Number: PCT/GB93/01651

(22) International Filing Date: 5 August 1993 (05.08.93)

(30) Priority data:
9216654.5 5 August 1992 (05.08.92) GB

(71) Applicant *(for all designated States except US)*: IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED [GB/GB]; Sardinia House, Sardinia Street, London WC2A 3NL (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)* : PARKER, Peter, Joseph, Jacques [GB/GB]; GOODE, Nigel, Thomas [AU/GB]; NURSE, Paul, Maxime [GB/GB]; Imperial Cancer Research Fund, P.O. Box 123, 44 Lincoln's Inn Fields, London WC2A 3PX (GB). WATERFIELD, Michael, Derek [GB/GB]; Ludwig Institute for Cancer Research, London W1P 8BT (GB).

(74) Agent: BASSETT, Richard, S.; Eric Potter Clarkson, St. Mary's Court, St. Mary's Gate, Nottingham NG1 1LE (GB).

(81) Designated States: GB, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: EUCARYOTIC CELLS TRANSFORMED WITH A MAMMALIAN PHOSPHOLIPID KINASE OR PROTEIN KINASE AND ASSAYS USING THEM

(57) Abstract

A eukaryotic cell transformed with a DNA construct comprising a coding sequence encoding a polypeptide having the activity of a mammalian phospholipid kinase or a mammalian protein kinase activated by a phospholipid or its metabolite and regulatory elements to allow transcription of the coding sequence in the said cell wherein the regulatory elements include a repressible or inducible promoter and the expression of the said coding sequence is lethal or growth inhibitory to the cell. In a preferred embodiment the cells are *Schizosaccharomyce pombe*. The cells are used as the basis of an assay for compounds involved in cell growth regulation. Such compounds can be used to treat cancers and the formation of blood vessel plaques.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**MR**	Mauritania
AU	Australia	**GA**	Gabon	**MW**	Malawi
BB	Barbados	**GB**	United Kingdom	**NE**	Niger
BE	Belgium	**GN**	Guinea	**NL**	Netherlands
BF	Burkina Faso	**GR**	Greece	**NO**	Norway
BG	Bulgaria	**HU**	Hungary	**NZ**	New Zealand
BJ	Benin	**IE**	Ireland	**PL**	Poland
BR	Brazil	**IT**	Italy	**PT**	Portugal
BY	Belarus	**JP**	Japan	**RO**	Romania
CA	Canada	**KP**	Democratic People's Republic of Korea	**RU**	Russian Federation
CF	Central African Republic			**SD**	Sudan
CG	Congo	**KR**	Republic of Korea	**SE**	Sweden
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovak Republic
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LU**	Luxembourg	**TD**	Chad
CS	Czechoslovakia	**LV**	Latvia	**TG**	Togo
CZ	Czech Republic	**MC**	Monaco	**UA**	Ukraine
DE	Germany	**MG**	Madagascar	**US**	United States of America
DK	Denmark	**ML**	Mali	**UZ**	Uzbekistan
ES	Spain	**MN**	Mongolia	**VN**	Viet Nam
FI	Finland				

EUCARYOTIC CELLS TRANSFORMED WITH A MAMMALIAN PHOSPHOLIPID KINASE OR PROTEIN KINASE AND ASSAYS USING THEM

The present invention relates to assays for compounds involved in cell growth regulation, and more particularly to those involved in transducing signals from hormones, growth factors and oncogenes. Such compounds represent potential drugs or targets for drugs to treat cancers and to prevent the formation of plaques which cause heart disease.

Phosphatidylinositol 3-OH kinase (PtdIns 3-kinase) catalyses the phosphorylation of the 3-hydroxyl of inositol in PtdIns, PtdIns-4-phosphate or in PtdIns-4,5-bisphosphate. This activity is involved in transducing signals from a number of hormones, growth factors and oncogenes. The standard assay for the activity of the PtdIns 3-kinase involving lipid moieties does not readily lend itself to a screen for potential inhibitors (or activators) of catalytic function. Members of the protein kinase C family of enzymes are involved in transducing signals from a number of hormones, growth factors and oncogenes. The standard assay for protein kinase C does not lend itself to a screen for potential inhibitors (or activators) of catalytic function.

Thus it has been desirable to investigate other means of searching for inhibitors.

A first aspect of the invention provides a eukaryotic cell transformed with a DNA construct comprising a coding sequence encoding a polypeptide having the activity of a mammalian phospholipid kinase or a mammalian protein kinase activated by a phospholipid or its metabolite and regulatory elements to allow transcription of the coding sequence in the said cell wherein the regulatory elements include a repressible or inducible promoter and the expression of the said coding sequence is lethal or growth inhibitory to the cell.

Polypeptides having the activity of a phospholipid kinase or a protein kinase activated by a phospholipid or its metabolite are involved in cell growth regulation.

By "growth inhibitory" we mean that the growth rate of cells transformed with the said DNA construct is at least two to three fold lower than the same cells not transformed with the said DNA construct when grown in the same culture conditions.

By "repressible" we mean that in the presence of a repressing agent the expression from the promoter is at least two-fold lower than expression from the promoter in the absence of the repressing agent.

It is preferred if expression from the promoter in the presence of a repressing agent is at least five-fold lower, more preferably ten-fold lower or even more preferably 100-fold lower than expression from the promoter in the absence of the repressing agent.

By "inducible" we mean that in the presence of an inducing agent the expression from the promoter is at least two-fold higher than expression from the promoter in the absence of the inducing agent.

It is preferred if expression from the promoter in the presence of an inducing agent is at least five-fold higher, more preferably ten-fold higher or even more preferably 100-fold higher than expression from the promoter in the absence of the inducing agent.

When an inducible promoter is used there is sufficiently low expression of the polypeptide in the uninduced state that the lethal or growth inhibitory phenotype is not observed whereas when the inducing agent is present the lethal or growth

inhibitory phenotype is observed.

When a repressible promoter is used there is sufficiently low expression of the polypeptide in the repressed state that the lethal or growth inhibitory phenotype is not observed whereas when the repressing agent is absent the lethal or growth inhibitory phenotype is observed.

Suitable eukaryotic cells include mammalian cells, such as COS cells and CHO cells, insect cells, slime mould such as *Dictyostelium*, and yeast.

Suitable regulatable mammalian cell promoters include glucocorticoid-inducible promoters and the metallothionein promoter.

It is preferred if the cell is a yeast cell.

Exemplary genera of yeast contemplated to be useful in the practice of the present invention are *Pichia, Saccharomyces, Kluyveromyces, Candida, Torulopsis, Hansenula, Schizosaccharomyces, Citeromyces, Pachysolen, Debaromyces, Metschunikowia, Rhodosporidium, Leucosporidium, Botryoascus, Sporidiobolus, Endomycopsis*, and the like.

It is preferred if the yeast is a fission yeast.

It is further preferred if the yeast is *Schizosaccharomyces*.

Preferably, the said polypeptide has the activity of a phospholipid kinase, for example a catalytically effective portion of the said kinase. Phospholipid kinases include phosphatidyl inositol 3-kinase, phosphatidyl inositol 4-kinase and phosphatidyl inositol 5-kinase which phosphorylate the inositol ring on the 3′, 4′ or 5′ hydroxyl, respectively.

Suitably, the said polypeptide is a catalytically effective portion of a phosphatidylinositol 3-OH kinase. It is convenient to use the 110 kDa mammalian PtdIns 3-kinase catalytic subunit.

In further preference, the said polypeptide is a catalytically effective portion of a protein kinase C (PKC). Suitably, the protein kinase C is PKC-γ or PKC-δ or PKC-η or PKC-ϵ.

A constitutive promoter such as *adh* may be used (disclosed in ref 1). Also, the SV40 promoter may be used.

Thus, a further aspect of the invention provides a *Schizosaccharomyces* cell transformed with a DNA construct comprising a coding sequence encoding a polypeptide having the activity of a mammalian phospholipid kinase or a mammalian protein kinase activated by a phospholipid or its metabolite and regulatory elements to allow transcription of the coding sequence in the said cell wherein the regulatory elements include a constitutive promoter and the expression of the said coding sequence is lethal or growth inhibitory to the cell.

Any gene that arrests growth or is lethal can be expressed only transiently for the purposes of subsequent inhibitor screening. In the case of a constitutive promoter in a plasmid carrying a marker, freshly transfected cells are diluted directly into medium using a combination of growth conditions to select for transfectants (for example, medium containing no leucine) and added potential inhibitors of the constitutively expressed mammalian gene to test for their efficacy.

Mammalian genes whose expression can be controlled by growth conditions can be introduced into the yeast under conditions where expression is low (ie suppressed or not induced).

It is preferred if the mammalian genes so introduced are stably maintained in the yeast.

It is further preferred if the mammalian genes are stably integrated into the yeast genome.

Expression is then increased following growth under de-repressing conditions (for example removal of thiamine) and potential inhibitors scored on their ability to permit growth under these conditions. The use of an integrant and a controllable promoter provides the most amenable procedure. The induction of cell arrest or cell death provides a powerful screen for a suppressor of such events. The present invention provides a screen for suppressors of regulatory proteins that control other mammalian functions either directly, for example protein kinases, or indirectly through the production of small regulatory molecules, for example an inositol lipid kinase.

Thus, in a preferred embodiment, the *S. pombe* cells contain a coding sequence for the 110 kDa mammalian PtdIns 3-kinase catalytic subunit under the regulatory control of the *nmt* promoter and with other suitable regulatory elements, such as a transcription terminator, as is known in the art, for expression of the said catalytic subunit. In the presence of thiamine the promoter is inoperative and the cells carrying the PtdIns 3-kinase catalytic subunit plasmid grow as the parental strain. (It will be appreciated by those skilled in the art that the parental strain may not be wild-type. For example mutant strains containing Ade$^-$ or Leu$^-$ or Ura$^-$ mutations may be used as the parental strain to allow selection of plasmid uptake). In the absence of thiamine the *nmt* promoter functions and the PtdIns 3-kinase catalytic subunit is induced. This has been shown by demonstrating a substantial increase in PtdIns 3-kinase activity under these conditions. However, following this induction the cells cease to divide; cultures plated in the absence of thiamine

do not grow but die.

Derivative of the *nmt* promoter that retain the thiamine-repressibility characteristics of the wild type promoter may also be used.

As an alternative to the thiamine-repressible *nmt* promoter, the *fbp1* gene promoter from *S. pombe* can be used. The *fbp1* gene promoter is repressed in the presence of 8% glucose as disclosed by Hoffman & Winston (1990) *Genetics* **124**, 807-816 incorporated herein by reference. Thus, in a further embodiment, the *S. pombe* cells contain a coding sequence for the 110 kDa mammalian PtdIns 3-kinase catalytic subunit under the regulatory control of the *fbp1* promoter and with other suitable regulatory elements for expression of the said catalytic subunit. In the presence of 8% glucose the function of the promoter is repressed and the cells carrying the PtdIns 3-kinase catalytic subunit plasmid grow on the parental strains. In the absence of glucose the *fbp1* promoter functions and the PtdIns 3-kinase catalytic subunit is induced.

The lethal phenotype of the *S. pombe* expressing mammalian PtdIns 3-kinase provides a very powerful tool with which to screen for inhibitors of this activity. Cells plated in the absence of thiamine will survive and proliferate if the activity of the PtdIns 3-kinase is suppressed. A direct demonstration that this is indeed the case, is afforded by the finding that a mammalian PtdIns 3-kinase regulatory subunit (p85α) when coexpressed with the PtdIns 3-kinase catalytic subunit will rescue these cells and allow proliferation. Clearly, therefore, coexpression of (or generally the presence of) the p85α subunit should be avoided in the assay of this embodiment, as should, in other embodiments, other activity-suppressing compounds.

In further embodiments the *S. pombe* cells contain a coding sequence for a mammalian protein kinase C under the regulatory control of the *nmt* promoter

or the *fbp1* promoter.

As an inhibitor screening process, a further advantage afforded by this approach is that general cytostatic and cytotoxic compounds will score negative; the screen will distinguish the action of the mammalian PtdIns 3-kinase or protein kinase C against the background of a plethora of essential eukaryotic gene functions.

Thus, a further aspect of the invention provides an assay kit comprising a eukaryotic cell according to the first aspect of the invention and culture medium such that the cell will divide and grow and such that the said coding sequence is expressed, the expressed polypeptide at least preventing cell division in the cell culture.

Conveniently the kit comprises *S. pombe* as the eukaryotic cell.

The invention also encompasses compounds identified as being useful in the assays of the invention.

These compounds are useful in the treatment of disease and medical conditions where there is an undesirable function of a phospholipid kinase or a protein kinase activated by a phospholipid or its metabolite.

Such diseases and conditions include cancer, inflammation, Alzheimer's disease, restenosis, atherosclerosis and wound healing.

Suitable promoters and coding sequence can be incorporated into vectors in the correct orientation by methods known in the art, some of which are described in Sambrook *et al* (1989) *Molecular Cloning, a practical approach* (2nd Edition), Sambrook, J., Fritsch, E. & Maniatis, T., eds, Cold Spring Harbor

Laboratory Press, Cold Spring Harbor, New York, incorporated herein by reference.

A variety of methods have been developed to operatively link DNA to vectors via complementary cohesive termini. For instance, complementary homopolymer tracts can be added to the DNA segment to be inserted to the vector DNA. The vector and DNA segment are then joined by hydrogen bonding between the complementary homopolymeric tails to form recombinant DNA molecules.

Synthetic linkers containing one or more restriction sites provide an alternative method of joining the DNA segment to vectors. The DNA segment, generated by endonuclease restriction digestion as described earlier, is treated with bacteriophage T4 DNA polymerase or *E. coli* DNA polymerase I, enzymes that remove protruding, 3'-single-stranded termini with their 3'-5'-exonucleolytic activities, and fill in recessed 3'-ends with their polymerizing activities.

The combination of these activities therefore generates blunt-ended DNA segments. The blunt-ended segments are then incubated with a large molar excess of linker molecules in the presence of an enzyme that is able to catalyze the ligation of blunt-ended DNA molecules, such as bacteriophage T4 DNA ligase. Thus, the products of the reaction are DNA segments carrying polymeric linker sequences at their ends. These DNA segments are then cleaved with the appropriate restriction enzyme and ligated to an expression vector that has been cleaved with an enzyme that produces termini compatible with those of the DNA segment.

Synthetic linkers containing a variety of restriction endonuclease sites are commercially available from a number of sources including International Biotechnologies Inc, New Haven, CN, USA.

A desirable way to modify the DNA encoding the polypeptide of the invention is to use the polymerase chain reaction as disclosed by Saiki *et al* (1988) *Science* 239, 487-491.

In this method the DNA to be enzymatically amplified is flanked by two specific oligonucleotide primers which themselves become incorporated into the amplified DNA. The said specific primers may contain restriction endonuclease recognition sites which can be used for cloning into expression vectors using methods known in the art.

Transformation of appropriate cell hosts is accomplished by well known methods that typically depend on the type of vector used and host cell. Transformation of *Saccharomyces* and related cells is described in Sherman *et al* (1986) *Methods In Yeast Genetics, A Laboratory Manual*, Cold Spring Harbor, NY. The method of Beggs (1978) *Nature* 275, 104-109 is also useful. With regard to vertebrate cells, reagents useful in transfecting such cells, for example calcium phosphate and DEAE-dextran or liposome formulations, are available from Stratagene Cloning Systems, or Life Technologies Inc., Gaithersburg, MD 20877, USA.

Schizosaccharomyces pombe may be transformed following LiCl treatment or by electroporation.

Conveniently, a Bio-Rad Pulse Controller may be used for electroporation of *S. pombe* cells.

a) Grow up cells to OD_{595} less than or equal to 0.5 in minimal medium.

b) Centrifuge cells at 1500 g for 5 min, remove supernatant and resuspend in 20 ml ice-cold distilled water, centrifuge again, remove supernatant and

resuspend in 20 ml ice-cold 1 M sorbitol, centrifuge again and remove supernatant.

c) Resuspend cells in ice-cold 1 M sorbitol to a density of $\sim 5x10^9$ cells/ml (concentrated 500 times when compared to original culture).

d) Use 40-100 μl of cell suspension per transformation. Add DNA (up to 100 ng) in 1 μl in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0) to cells and incubate on ice 5 min.

e) Transfer to pre-chilled cuvettes (0.2 cm gap) and apply pulse (1.5 KV, 25 μF, 200 Ω).

f) Immediately add 900 μl of ice-cold 1 M sorbitol and transfer to a chilled tube on ice.

g) Promptly spread 100-200 μl onto a selective minimal medium plate containing 1 M sorbitol and culture at 32°C until grown.

The technique of electroporation of yeast is disclosed in Becker, D.M. and Guarente, L. (1990) *Meth. Enzymol.* **194**, 182.

Machines for electroporation are available from other manufacturers and can be used to transform yeast and mammalian cells according to their instructions.

Successfully transformed cells, ie cells that contain a DNA construct, can be identified by well known techniques. For example, cells resulting from the introduction of an expression construct of the present invention can be grown to produce the polypeptide of the invention. Cells can be harvested and lysed and their DNA content examined for the presence of the DNA using a method

such as that described by Southern (1975) *J. Mol. Biol.* 98, 503 or Berent *et al* (1985) *Biotech.* 3, 208.

In addition to directly assaying for the presence of recombinant DNA, successful transformation can be confirmed by well known immunological methods when the recombinant DNA is capable of directing the expression of the protein. For example, cells successfully transformed with an expression vector produce proteins displaying appropriate antigenicity. Samples of cells suspected of being transformed are harvested and assayed for the protein using suitable antibodies, for example by western blotting.

The invention will now be described in detail with reference to the following Examples and Figures wherein:

Figure 1 shows the nucleotide sequence (SEQ ID No 1) and deduced amino acid (SEQ ID No 2) of the sequence 110 kDa catalytic subunit of PtdIns 3-kinase (P110).

Figure 2 shows the nucleotide sequence of the *nmt* promoter region (SEQ ID No 3).

Figure 3 shows the nucleotide sequence of PKC-ϵ (SEQ ID No 4).

Figure 4 shows the nucleotide sequence of PKC-γ (SEQ ID No 5).

Figure 5 shows the nucleotide sequence of PKC-δ (SEQ ID No 6).

Figure 6 shows the nucleotide sequence of PKC-η (SEQ ID No 7).

Figure 7 shows that the lethal effect of p110 expression in *S. pombe* is

suppressed by p85 expression.

Figure 8 shows the isotype-specific effects of PKC expression in *S. pombe*.

Figure 9 shows the effect of PKC expression on growth rates in liquid culture.

Figure 10 shows that PKC-δ-induced growth inhibition is the result of kinase activity.

Example 1: Assay using catalytic subunit of PtdIns 3-kinase and *nmt* promoter

Isolation of PtdIns 3-kinase catalytic subunit cDNA. The cDNA for the 110 kDa catalytic subunit can be isolated by a conventional cloning strategy. Purification of the bovine enzyme from brain tissue (Morgan, Smith *et al* 1990) has demonstrated that sufficient protein can be isolated for protein sequence determination. This is unequivocally established for the 85 kDa regulatory subunit which has been sequenced from this source and, as a consequence, cloned (Otsu, Hiles *et al* 1991). The PtdIns 3-kinase from bovine brain (85-110 dimer) is purified according to Morgan, Smith *et al* (1990) by sequential fractionation with ammonium sulphate and chromatography on DEAE-cellulose, phosphocellulose, Sephacryl S-200 and Mono Q. In order to remove contaminants and separate subunits, the protein is further purified by sodium dodecyl sulphate polyacrylamide gel electrophoresis according to Laemmli (1970), the 110 kDa protein visualised in ammonium chloride (4N), electroeluted and digested with trypsin as described in Katan, Kriz *et al* (1988). Tryptic peptides are then separated by standard procedures and subjected to amino acid sequence determination. Sequence established for the 110 kDa catalytic subunit is used to predict redundant oligonucleotide probes for screening a bovine brain cDNA library. Standard cloning procedures are then

employed in the isolation of a cDNA encoding the complete open reading frame of the 110 kDa subunit (Sambrook *et al* 1989). The sequence of the cDNA is determined by commonly employed dideoxy-sequencing procedures. A specific example of using this strategy is described by Hiles *et al* (1992) *Cell* 70, 419-429.

Materials: Restriction enzymes and DNA modification enzymes were obtained from standard commercial sources and used according to the manufacturer's recommendations. Oligonucleotides were synthesised on an Applied Biosystems 380B DNA synthesiser and used directly in subsequent procedures.

Protein Purification and Amino Acid Sequence Determination: The purification of the p85α and p110 proteins by chromatography on a peptide affinity column corresponding to amino acids 742-758 of the kinase insert region of the human PDGF-β receptor has been described (Otsu *et al* (1991) *Cell* **65**, 91-104). Proteins were released from the affinity matrix using SDS-containing buffers, separated on a Prosieve agarose gel, and visualised by staining with Coomassie blue G250. The band corresponding to p110 was excised and protein was eluted by tube gel HPEC. Protein was precipitated from p110-containing fractions by treatment with trichloroacetic acid and then washed with acetone. The p110-containing pellet was resuspended and digested with lysylendopeptidase in the presence of SDS, and peptides were separated by tandem ion-exchange chromatography and reverse-phase HPLC. This procedure was carried out on three separate PI3-kinase preparations. A fourth preparation was eluted from the matrix as before and boiled for 5 min. After cooling, the sample was diluted with 25 mM Tris-HCl (pH 8.8) and digested directly with lysylendopeptidase for 72 hr at 30°C. Peptides were separated as above. Peptide sequences were determined using a modified Applied Biosystems 477A automated pulse-liquid sequencer.

mRNA Isolation and cDNA Cloning: Total RNA was isolated from SGBAF-1 cells by the method of Chirgwin *et al* (1979) *Biochemistry* 18, 294-299 and poly(A) mRNA was selected by chromatography on oligo-(dT)-cellulose (Maniatis *et al* (1982) *Molecular Cloning: A laboratory manual*, Cold Spring Harbor Press, Cold Spring Harbor, New York). An oligo-dT primed cDNA library of 5 x 10^6 primary recombinants was constructed in lambda Uni-Zap (Stratagene) from 5 μg of this mRNA using the Stratagene Uni-Zap cDNA cloning system. The construction of the total bovine brain cDNA library in lambda Uni-Zap has been described previously (Otsu *et al* (1991) *Cell* 65, 91-104).

Library Screening and Hybridisations: The unamplified SGBAF-1 cDNA library (10^6 recombinants) was plated on *E. coli* K12 PLK-F (Stratagene) at a density of 10^5 plaques per 15 cm dish, and lifts were taken in duplicate onto nitrocellulose membranes (Millipore). For screening, filters were prehybridised for at least 1 hr at 42°C in 6 x SSPE, 0.5% SDS, 10 x Denhardt's solution, and 100 μg/ml denatured sonicated herring sperm DNA (Sigma). Hybridisation was carried out in the same solution containing 10 ng/ml radiolabeled oligonucleotide. Oligonucleotides used were: peptide N, (MDWIFHT; SEQ ID No 8) 5'-AA(G/A)ATGGA(T/C)TGGAT(C/T/A)TT(T/C)CA(T/C)AC-3' (SEQ ID No 9); peptide J (DDGQLFHIDFGHF; SEQ ID No 10) 5'-GATGATGGCC-A(G/A)CTGTT(T/C)CA(T/C)AT(T/A)GA(T/C)-TTTGGCCA(T/C)TT (SEQ ID No 11). Oligonucleotides were labeled with ^{32}P at the 5' end in a 20 μl reaction containing 100 ng of oligonucleotide, 1 x kinase buffer (Promega), 0.1 mM spermidine, 5 mM dithiothreitol, 100 μCi of [γ-^{32}P]ATP (5000 Ci/mmol, Amersham), and 2 μl (20 U) of T4 polynucleotide kinase (Amersham). Filters were washed in 6 x SSC, 0.1% SDS at room temperature and then subjected to autoradiography using Kodak XAR film. Hybridising clones were plaque purified and rescued as plasmids according to the manufacturer's instructions.

Characterisation of cDNA Clones: Sequencing was carried out by the chain termination method using the Sequenase system (US Biochemicals). Clones for sequencing were obtained by directed cloning of restriction fragments into M13mp18 and mp19 vectors (Yanisch-Perron *et al* (1985) *Gene* 33, 103-119) and by making a series of exonuclease III-mediated deletions (Henikoff (1984) *Gene* 28, 351-359; Pharmacia Exonuclease III deletion kit). DNA sequences were analysed on a Micro-VAX computer using the Wisconsin sequence analysis package (UWGCG; Devereux *et al* (1984) *Nucl. Acids Res.* 12, 387-395).

RACE PCR: RACE PCR was carried out essentially as described previously (Frohman *et al* (1988) *Proc. Natl. Acad. Sci. USA* 85, 8998-9002; Harvey and Garlison (1991) *Nucl. Acids Res.* 19, 4002). In brief, first-strand cDNA primed with random hexamers (Amersham) was synthesised from 1 μg of SGBAF-1 cell mRNA using the Stratagene first-strand cDNA synthesis kit. First-strand cDNA was isolated by isopropanol precipitation and tailed with oligo-(dA) using terminal deoxynucleotidyl transferase (Bethesda Research Laboratories). PCR was performed using oligo 2224 (5'-AATTCACACACTGGCATGCCGAT; SEQ ID No 12) and adaptor dT (5'-GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT; SEQ ID No 13) as primers, using a Perkin-Elmer Cetus Taq polymerase PCR kit (conditions: 30 cycles of 94°C for 1 min, 35°C for 1 min, 72°C for 2 min). Products were fractionated on a 1.5% low melting point agarose gel and visualised by staining with ethidium bromide. The gel was sliced into six bands (ranging from 150 bp to 2000 bp), and DNA was isolated from each gel slice. A further round of PCR was performed on this DNA using oligonucleotide 2280 (5'-TTTAAGCTTAGGCATTCTAAAGTCACTATCATCCC; SEQ ID No 14) and adaptor (5'-GACTCGAGTCGACATCGA; SEQ ID No 15) as primers (conditions: 35 cycles of 94°C for 1 min, 56°C for 1 min, 72°C for 2 min). Products were fractionated on an agarose gel and visualised by staining with

ethidium bromide. A band 250 bp shorter than the size of the DNA in the gel slice used for the PCR was expected. An intensely staining band of 350 bp obtained from the ~600 bp gel slice was excised, digested with *Hind*III and *Sal*I, and ligated into Bluescript KS- digested with *Hind*III and *Xho*I to give plasmid pBS/race. Two independent inserts were completely sequenced. The sequence of p110, the 110-kD catalytic subunit of PI3-kinase is shown in Figure 1 and has the GenEMBL Accession No M93259 (SEQ ID No 1).

Isolation of *nmt* promoter. The promoter has been isolated by Maundrell (2) and may be isolated by repeating the procedures reported in that reference. Moreover, the sequence of the gene, including the promoter, has been submitted to the GenBank™/EMBL database as Accession No J05493 and is shown in Figure 2 (SEQ ID No 3).

Vectors containing the *nmt* promoter and derivatives of the *nmt* promoter suitable for use in the present invention are described by Basi *et al* (1992) *Yeast* 8, S597 (special issue) and Maundrell (1990).

The upstream regulatory region and downstream polyadenylation site of *nmt1* have been incorporated into two types of *S. pombe/E. coli* shuttle vector: pREP extrachromosomally replicating plasmids and pRIP integrating plasmids. Using either of these constructs thiamine mediated transcriptional regulation can be transferred to heterologous coding sequences.

The time course of induction and repression have been studied as a function of changes in the intracellular thiamine concentration. Addition of thiamine to cells growing in minimal medium results in a rapid rise in the internal thiamine from a basal level of around 10 pmoles/10^7 cells to up to 1000 fold this level and this is accompanied by repression of *nmt1* promoter activity. If cells are then washed and allowed to continue growth in minimal medium, the

intracellular thiamine is progressively diluted as the cell mass doubles and transcription is reinitiated as the internal thiamine concentration falls below 50 pmoles/10^7 cells. The time taken to re-activate the *nmt1* promoter therefore depends on the internal thiamine concentration at the time when the cells are transferred to thiamine free medium.

Quantitation of promoter strength was assessed using chloramphenicol acetyl transferase as a reporter gene. The fully induced *nmt1* promoter is about 6 fold more active than the *S. pombe adh* promoter and its activity is reduced about 80 fold when cells are grown in repressing conditions. These vectors are ideally suited to applications requiring maximal expression of a gene of interest. In addition, two modified versions of the promoter with reduced activity have been created following an analysis of the effects of TATA box mutations. Truncating the wild type TATA box, TATATAAA to ATAAA (the '4' series) or AT (the '8' series) down-regulates transcriptional activity of the *nmt1* promoter by approximately 1 and 2 orders of magnitude respectively (see Table). These mutations in the TATA box do not affect thiamine repressibility or the site of transcription initiation.

The table below summarises the salient features of some of the vectors which have been constructed thus far:

vector	TATA box	selectable marker	restriction site at ATG	relative activity[a] -thiamine	+thiamine
pREP1	TATATAAA	LEU2	*Nde*I	80	1
pREP2	TATATAAA	ura4	*Nde*I	80	1
pREP3	TATATAAA	LEU2	*Bal*I	80	1
pREP3X	TATATAA	LEU2	-[b]	80	1
pREP4	TATATAAA	ura4	*Bal*I	80	1
pREP41	ATAAA	LEU2	*Nde*I	12	0.06
pREP42	ATAAA	ura4	*Nde*I	12	0.06
pREP6	TATATAAA	Sup3.5	*Bal*I	80	1
pREP6X	TATATAAA	Sup3.5	multiple cloning site	80	1
pREP81	AT	LEU2	*Nde*I	1	0.004
pREP82	AT	ura4	*Nde*I	1	0.004

[a] activity is based on the quantitation of CAT assays. Data are expressed in arbitrary units relative to the wild type promoter cultured in the presence of thiamine.

b the *Ball* site is replaced with an *Xho*I site allowing expression from the ATG.

c in some of the vectors the complementation gene used for selection of plasmid uptake has been changed from the LEU2 gene to the sup 3.5 gene which complements the Ade 6.704 mutation or to the URA4 gene. The backbone of the plasmid is not altered (ie promoter and stop sequence from the *nmt*1 gene, ARS1 and pUC119 backbone).

Construction of an *S. pombe* p110 expression system. A suitable restriction fragment containing the complete 110 kDa subunit open reading frame and flanking sequences is subcloned into the *nmt* promoter plasmid containing a suitable marker gene for selection creating an *nmt*-100 plasmid in order to allow expression of the 110 kDa protein under the control of the thiamine repressible *nmt* promoter. The *nmt*-110 plasmid is grown in a suitable bacterial host and the plasmid purified by conventional techniques (Sambrook *et al* 1989). A 3.4 kb *Bam*HI/*Fsp*I fragment containing the cDNA of p110 was isolated and subcloned into the *Bam*HI/*Sma*I sites of pREP3X-p110 (*nmt*-110).

The *nmt*-110 plasmid is then transfected by standard procedures (Moreno, Klar *et al* 1991) into a *Schizosaccharomyces pombe* strain that is auxotrophic for leucine cells are transformed using electroporation. Transfected cells are then plated in the presence of thiamine and in the absence of leucine. As an alternative *Schizosaccharomyces pombe* strains which are auxotrophic for adenine or uracil (that is Ade^- or Ura^-) may be used; in this case the cells are plated in the presence of thiamine and absence of adenine, or the presence of thiamine and absence of uracil, respectively. Colonies growing up under these conditions are then analysed for the presence of the *nmt*-110 plasmid. The lethal phenotype caused by the expression of 110 kDa protein is checked by replating colonies in the presence or absence of thiamine; under the latter conditions colonies will arrest and/or die.

For the purposes of setting up a screen for inhibitors, a stable transformant is isolated. This is carried out by standard procedures involving growth in the presence and absence of the selectable marker leucine (or adenine or uracil). Isolates obtained in this manner are checked for the stable insertion of the 110 kDa sequence into genomic DNA by Southern analysis or stable replication of a non-integrated plasmid. Expression of the p110 protein is also confirmed by western blot analysis of the transformants using antibodies reactive against p110, or by measuring the activity of the p110 subunit in the transformed cells. The inducible lethal phenotype is rechecked by growth of these isolates in the presence and absence of thiamine ($\geq$ 10 nM).

It is preferred if 100 nM, or > 1 pM or > 1 μM is used.

It is most preferred if 15 μM thiamine is used.

Operating the screen. The screen for inhibitor activity is carried out on a 96-well microtitre plate format. An integrant colony is picked and put into liquid culture in minimal medium, 2% glucose, 15 μM thiamine and supplements appropriate for the strain (eg uracil 50 μg/ml would be included for a ura$^-$ strain if the integrated plasmid did not harbour a URA4-based selection marker). This culture is grown up and, after extensive washing, used to seed two 10 ml cultures, one containing thiamine as above, and one without. The cultures are expanded overnight and then diluted to an optical density (OD) at 595 nm of 0.01-0.10. For those cells requiring treatment for arrest of growth additions are made at this stage prior to plating. The diluted cultures are then aliquoted into wells of a sterile 96 well microtitre plate containing individual test compounds in the presence or absence of thiamine. The growth of the cells is monitored over time until the OD_{595} reached is ~0.8 for control cultures. Control cultures are those cultured with thiamine. The OD_{595} is assessed using a microtitre plate reader.

The cells precultured in thiamine and retained in thiamine serve to indicate optimum growth rate. Cells precultured in the absence of thiamine and then put into wells containing thiamine provide a control for the rescue of growth. Cells precultured in the absence of thiamine and put into wells in the absence of thiamine or test compound provide a baseline for non-growth. Individual test compounds are assessed for their potency in permitting growth in the absence of thiamine in cells plated in the absence of thiamine.

Accumulated experience in the operation of this screen for a particular gene product permits a less frequent monitoring of the growth curves and a single time point may be found to be sufficient. Similarly, cultures propagated throughout in the presence of thiamine may be found to be a non-essential control. These alterations to the procedure may provide some practical advantages in increasing the number of test compounds per 96 well plate and in reducing the time required for assessment of growth.

The above procedures have been employed in creating an *S. pombe* strain harbouring a p110 cDNA under the control of the *nmt* promoter. Switching these cells from a medium containing thiamine (15 μM) to one in the absence of thiamine causes growth arrest. Evidence that the arrest is a consequence of the expression of the mammalian protein has come from a number of observations:

1. Transient transfection and subsequent expression has been observed on multiple occasions with the p110 cDNA and not with the vector alone.

2. On expression of the p110 protein, it is possible to detect the activity of the expressed mammalian protein in cell extracts, ie the catalytic activity is retained on expression in *S. pombe*.

3. On expression of the mammalian regulatory subunit of the kinase, p85α [4], increased expression of p110 no longer induces growth arrest.

The use of this system as a viable tool for screening p110 inhibitors is evidenced by the ability of p85α, the regulatory subunit, to suppress the growth arrest phenotype. Biochemical evidence has already established that the p85α-p110 complex is less active than the free p110 protein [9].

The lethal effect of p110 expression in *S. pombe* is suppressed by p85 expression as shown in Figure 7. Stable p110-expressing *S. pombe* cells were transformed with the pREP4 vector, or the pREP4-p85α or pREP4-p85β constructs and, after selection for plasmid uptake, were streaked onto selective minimal medium plates in the presence or absence of thiamine. Expression of p110 alone is lethal but this effect is rescued by co-expression of either p85α or p85β.

The p85α and p85β cDNAs can be obtained using the methods described by Otsu *et al* (1991) *Cell* **65**, 91-104 incorporated herein by reference.

Example 2: Isotype-specific effects of PKC expression in *S. pombe* and the effect of PKC expression on growth rates in liquid culture

S. pombe strains containing integrated plasmids for expression of mammalian PKC-γ, -δ, -ε, -ζ or -η were streaked onto selective minimal medium plates in the absence of thiamine or the presence of thiamine or TPA as shown in Figure 8. Growth of control (vector) or PKC-ζ cells was similar under all three conditions. PKC-γ expression (Figure 8, plate B) marginally decreased growth and TPA addition to these cells totally supressed growth (Figure 8, plate C). PKC-δ, -ε and -η expression alone was markedly growth inhibitory (Figure 8, plate B).

Stable PKC-*S. pombe* strains were cultured in minimal medium in the absence of thiamine for 18 hours until an OD^{595} of 0.2-0.5 was attained (see Figure 9). Strains were then (at time zero) diluted to an OD^{595} of 0.02 in minimal medium and cultured in the presence of 1 μM thiamine (controls) (▲), in the absence of thiamine (■) or in the absence of thiamine with 100 ng/ml TPA (○). At the indicated times, the cell density was calculated by measuring the OD^{595}. PKC-ζ cells grew at a rate essentially indistinguishable from vector controls. PKC-δ, -ϵ and -η expression markedly delayed growth when compared with vector controls (-thiamine). Growth of PKC-γ, -δ and -η expressing cells was essentially nil when cultured in the presence of TPA.

Example 3: An inhibitor screen for protein kinase C-ϵ

Protein kinase C-ϵ [10] cDNA (Figure 3; SEQ ID No 4) has been introduced into a plasmid under the control of the *nmt* promoter yielding *nmt*-PKC-ϵ. A 2.7 kb *Xho*I fragment with the full coding sequence for PKC-ϵ was isolated from pMT2-PKC-ϵ and subcloned into *Sal*I-digested pREP3X. Then 300 bp of 5′ non-coding sequence was removed by digesting with *Xho*I and *Nco*I, blunting the ends and religating to give pREP3X-PKC-ϵ. The plasmid pMT2-PKC-ϵ can be prepared by the methods described by Schaap *et al* (1989) ***FEBS Lett.*** **243**, 351-357. Transfection of this construct into *S. pombe* employing selection for uptake of the LEU2 gene in the presence of thiamine, yields populations of cells that on switching to "no thiamine" conditions while retaining selection for LEU2, reduce growth rate.

Growth inhibition is consistent with the expression of the mammalian PKC-ϵ gene product since:

1. Growth inhibition correlates with an induction of the PKC-ϵ protein as

judged by Western analysis.

2. The induced phenotype also correlates with expression of PKC-ϵ activity as determined in cell extracts.

3. Suppression of PKC-ϵ expression by exposure to the phorbol ester TPA can rescue cells that are expressing low levels of PKC-ϵ (cells expressing high levels of PKC-ϵ are not rescued and the steady state level of PKC-ϵ is not significantly depressed by TPA treatment).

The expression of a functional PKC-ϵ activity in *S. pombe* and its correlation with growth arrest under various growth conditions provides the basis for an inhibitor screen.

The transformed cells are plated in the presence of thiamine (control) and the absence of thiamine (test) and the compound to be assayed is added to the "test" plates.

Example 4: An inhibitor screen for protein kinase C-γ.

A cDNA for PKC-γ (Figure 4; SEQ ID No 5) has been introduced into a plasmid under the control of the *nmt* promoter, producing *nmt*-PKC-γ. A 2.4 kb *Bam*HI/blunt *Hin*dIII fragment with the full coding sequence of PKC-γ was isolated from pSP64-PKC-γ and subcloned into the *Bam*HI/*Sma*I sites of pREP3X to give pREP3X-PKC-γ. The plasmid pSP64-PKC-γ can be prepared as described by Patel & Stabel (1989) *Cell. Signall.* **1**, 227-240. Transfection of *S. pombe* with *nmt*-PKC-γ yields populations of cells that on switching to medium without thiamine induce PKC-γ protein as determined by Western blotting and the detection of PKC activity in cell extracts. These cells continue to grow on induction but if the PKC-γ is selectively activated by inclusion of

the phorbol ester TPA in the growth medium, the cells will arrest. The dependence of growth arrest upon the inclusion of TPA provides direct evidence that the catalytic function of PKC-γ is responsible for the phenotype. No such arrest is observed on treatment of the original *S. pombe* strain. Other PKC activators, such as Mezerein, or other phorbol esters or diacylglycerols may be used in place of TPA.

That activation of PKC-γ induces growth arrest provides a screen for inhibition of function of this mammalian gene product.

Operating the screen. The screen for inhibitor activities is carried out on a 96-well microtitre plate format. For thiamine repressible genes, stable integrants are grown up overnight (12 h) in the absence of thiamine. The culture is then diluted in the absence of thiamine to an OD_{595} = 0.01 to 0.10. The culture is then aliquoted into microtitre wells containing the potential inhibitors and, in the case of PKC-γ, also phorbol ester. The growth of cells monitored at 595 nm using a microtitre plate reader. Cells are allowed to grow until parallel wells/plates containing cells growing in the presence of thiamine (15 μM) have increased their OD_{595} to 1.0 units. Cells from the test wells that have proliferated can be scored relative to both control wells (eg +thiamine) and no addition wells (-inhibitor, -thiamine).

Thus, for PKC-γ there are the following possibilities: (i) control plates which are +thiamine or -thiamine or -thiamine + TPA and (ii) test plates which are +thiamine + compound or -thiamine + compound or -thiamine + TPA + compound.

Example 5: An inhibitor screen for protein kinase C-δ.

A cDNA for PKC-δ (Figure 5; SEQ ID No 6) has been introduced into a plasmid under the control of the *nmt* promoter, producing *nmt*-PKC-δ. A 2.4 kb blunt *Pfi*MI/*Nde*I fragment containing the full coding sequence of PKC-δ was isolated from pBluescript-PKC-δ and subcloned into blunt *Sal*I-digested pREP3X to give pREP3X-PKC-δ. The plasmid pBluescript-PKC-δ can be obtained using the methods described in Olivier & Parker (1991) *Eur. J. Biochem.* 200, 805-810 incorporated herein by reference. Transfection of *S. pombe* with *nmt*-PKC-δ yields populations of cells that on switching to medium without thiamine induce PKC-δ protein as determined by Western blotting and by activity measurements. There is marked growth inhibition by expression alone and if the PKC-δ is activated by inclusion of the phorbol ester TPA in the growth medium, the phenotype is strengthened. Experiments with PKC-δ also provide firm evidence that the phenotype is a result of the function of the kinase. Part of the kinase domain of PKC-δ was deleted thus rendering it enzymatically inactive. The product was expressed to a high level in *S. pombe* but there was no growth inhibition thus indicating that the phenotype is due to the functional kinase.

That activation of PKC-δ induces growth inhibition provides a screen for inhibition of function of this mammalian gene product.

Operating the screen. The screen for inhibitor activities is carried out on a 96-well microtitre plate format. For thiamine repressible genes, stable integrants are grown up overnight (12 h) in the absence of thiamine. The culture is then diluted in the absence of thiamine to an OD_{595} = 0.01 to 0.10. The culture is then aliquoted into microtitre wells containing the potential inhibitors and, in the case of PKC-γ, also phorbol ester. The growth of cells monitored at 595 nm using a microtitre plate reader. Cells are allowed to grow

until parallel wells/plates containing cells growing in the presence of thiamine (15 μM) have increased their OD_{595} to 1.0 units. Cells from the test wells that have proliferated can be scored relative to both control wells (+thiamine) and no addition wells (-inhibitor, -thiamine). Additionally, the test wells may contain or lack TPA.

Figure 10 shows that the PKC-δ-induced growth inhibition is the result of kinase activity. *S. pombe* cells were transformed with a control vector or vectors to express the full length PKC-δ protein or a PKC-δ protein in which part of the catalytic domain has been deleted to render it functionally inactive as a protein kinase (PKC-$\delta\Delta$). After selection for uptake of plasmid, a number of colonies were plated onto selective medium plates in the presence of thiamine, the absence of thiamine or the presence of TPA. PKC-δ expression markedly inhibits growth (-thiamine plate) and addition of TPA increases the effect. In contrast, expression of PKC-$\delta\Delta$ has no effect on growth under any condition.

Example 6: An inhibitor screen for protein kinase C-η.

A cDNA for PKC-η (Figure 6; SEQ ID No 7) has been introduced into a plasmid under the control of the *nmt* promoter, producing *nmt*-PKC-η. A 3.3 kb *Xho*I fragment containing the coding sequence for PKC-η was isolated from pBluescript-PKC-η and subcloned into *Sal*I-digested pREP3X to give pREP3X-PKC-η. The plasmid pBluescript-PKC-η can be obtained using the methods described by Dekker *et al* (1992) *FEBS Lett.* **312**, 195-199. Transfection of *S. pombe* with *nmt*-PKC-η yields populations of cells that on switching to medium without thiamine induce PKC-η protein as determined by Western blotting and the detection of PKC activity in cell extracts. However, there is some expression even in the presence of thiamine which produces ~50% growth inhibition. There is an even more marked growth inhibition by derepressed

expression alone and if the PKC-η is selectively activated by inclusion of the phorbol ester TPA in the growth medium, there is no growth.

That activation of PKC-η induces growth inhibition provides a screen for inhibition of function of this mammalian gene product.

Operating the screen. The screen for inhibitor activities is carried out on a 96-well microtitre plate format. For thiamine repressible genes, stable integrants are grown up overnight (12 h) in the absence of thiamine. The culture is then diluted in the absence of thiamine to an OD_{595} = 0.01 to 0.10. The culture is then aliquoted into microtitre wells containing the potential inhibitors and, in the case of PKC-γ, also phorbol ester. The growth of cells monitored at 595 nm using a microtitre plate reader. Cells are allowed to grow until parallel wells/plates containing cells growing in the presence of thiamine (15 μM) have increased their OD_{595} to 1.0 units. Cells from the test wells that have proliferated can be scored relative to both control wells (+thiamine) and no addition wells (-inhibitor, -thiamine). Additionally, the test wells may contain or lack TPA.

REFERENCES

1. Russell, P. & Nurse, P. (1986) *Cell* **45**, 145-153.

2. Maundrell, K. (1990) *J. Biol. Chem.* **265**, 10857-10864.

3. Morgan, S.J. *et al* (1990) *Eur. J. Biochem.* **191**, 761-767.

4. Otsu, M. *et al* (1991) *Cell* **65**, 91-104.

5. Laemmli, U.K. (1970) *Nature* **227**, 680-685.

6. Katan, M. *et al* (1988) *Cell* **54**, 171-177.

7. Sambrook, J. *et al* (1989) *"Molecular Cloning, a Laboratory Manual"* C.S.H. Laboratory Press.

8. Moreno, S. *et al* (1991) *"Molecular Genetic Analysis of fission yeast Schizosaccharomyces pombe"*. In *Methods in Enzymology*. pp 795-823.

9. Shibasaki, F. *et al* (1991) *J. Biol. Chem.* **266**, 8108-8114.

10. Schaap, D. *et al* (1989) *FEBS Lett.* **243**, 351-357.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Imperial Cancer Research Technology Ltd
(B) STREET: Sardinia House, Sardinia Street
(C) CITY: London
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): WC2A 3NL
(G) TELEPHONE: 071 242 1136
(H) TELEFAX: 071 831 4991

(ii) TITLE OF INVENTION: Transformed cells and assays using them

(iii) NUMBER OF SEQUENCES: 15

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3498 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..3204

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
ATG CCT CCA AGA CCA TCA TCA GGT GAA CTG TGG GGC ATC CAC TTG ATG          48
Met Pro Pro Arg Pro Ser Ser Gly Glu Leu Trp Gly Ile His Leu Met
 1               5                  10                  15

CCC CCA AGA ATC CTA GTA GAA TGT TTA CTA CCA AAT GGG ATG ATA GTG          96
Pro Pro Arg Ile Leu Val Glu Cys Leu Leu Pro Asn Gly Met Ile Val
            20                  25                  30

ACT TTA GAA TGC CTC CGT GAG GCT ACG TTA ATA ACG ATA AAG CAT GAA         144
Thr Leu Glu Cys Leu Arg Glu Ala Thr Leu Ile Thr Ile Lys His Glu
        35                  40                  45

CTA TTT AAA GAA GCA AGA AAA TAC CCT CTC CAT CAA CTT CTT CAA GAT         192
Leu Phe Lys Glu Ala Arg Lys Tyr Pro Leu His Gln Leu Leu Gln Asp
    50                  55                  60

GAA TCT TCT TAC ATT TTC GTA AGT GTT ACC CAA GAA GCA GAA AGG GAA         240
Glu Ser Ser Tyr Ile Phe Val Ser Val Thr Gln Glu Ala Glu Arg Glu
```

```
 65                  70                  75                  80

GAA TTT TTT GAT GAA ACA AGA CGA CTT TGT GAC CTT CGG CTT TTT CAA        288
Glu Phe Phe Asp Glu Thr Arg Arg Leu Cys Asp Leu Arg Leu Phe Gln
                    85                  90                  95

CCC TTT TTA AAA GTA ATT GAA CCA GTA GGC AAC CGT GAA GAA AAG ATC        336
Pro Phe Leu Lys Val Ile Glu Pro Val Gly Asn Arg Glu Glu Lys Ile
               100                 105                 110

CTC AAT CGA GAA ATT GGT TTT GCT ATC GGC ATG CCA GTG TGT GAA TTC        384
Leu Asn Arg Glu Ile Gly Phe Ala Ile Gly Met Pro Val Cys Glu Phe
           115                 120                 125

GAT ATG GTT AAA GAT CCA GAA GTA CAG GAC TTC CGA AGA AAT ATT CTC        432
Asp Met Val Lys Asp Pro Glu Val Gln Asp Phe Arg Arg Asn Ile Leu
       130                 135                 140

AAT GTT TGT AAA GAA GCT GTG GAT CTT AGG GAT CTT AAT TCA CCT CAT        480
Asn Val Cys Lys Glu Ala Val Asp Leu Arg Asp Leu Asn Ser Pro His
145                 150                 155                 160

AGT AGA GCA ATG TAT GTT TAT CCT CCA AAT GTA GAA TCT TCA CCA GAA        528
Ser Arg Ala Met Tyr Val Tyr Pro Pro Asn Val Glu Ser Ser Pro Glu
                   165                 170                 175

CTG CCA AAG CAC ATA TAT AAT AAA TTG GAT AAA GGG CAA ATA ATA GTG        576
Leu Pro Lys His Ile Tyr Asn Lys Leu Asp Lys Gly Gln Ile Ile Val
               180                 185                 190

GTG ATT TGG GTA ATA GTT TCT CCA AAT AAT GAC AAA CAG AAG TAT ACT        624
Val Ile Trp Val Ile Val Ser Pro Asn Asn Asp Lys Gln Lys Tyr Thr
           195                 200                 205

CTG AAA ATC AAC CAT GAC TGT GTG CCA GAA CAA GTA ATT GCT GAA GCA        672
Leu Lys Ile Asn His Asp Cys Val Pro Glu Gln Val Ile Ala Glu Ala
       210                 215                 220

ATC AGG AAA AAA ACT CGA AGT ATG TTG CTA TCA TCT GAA CAA CTA AAA        720
Ile Arg Lys Lys Thr Arg Ser Met Leu Leu Ser Ser Glu Gln Leu Lys
225                 230                 235                 240

CTC TGT GTT TTA GAA TAT CAG GGC AAG TAT ATT TTA AAA GTG TGT GGA        768
Leu Cys Val Leu Glu Tyr Gln Gly Lys Tyr Ile Leu Lys Val Cys Gly
                   245                 250                 255

TGT GAT GAA TAC TTC CTA GAA AAA TAT CCT CTG AGT CAG TAT AAG TAT        816
Cys Asp Glu Tyr Phe Leu Glu Lys Tyr Pro Leu Ser Gln Tyr Lys Tyr
               260                 265                 270

ATA AGA AGC TGT ATA ATG CTT GGG AGG ATG CCC AAT TTG ATG CTG ATG        864
Ile Arg Ser Cys Ile Met Leu Gly Arg Met Pro Asn Leu Met Leu Met
           275                 280                 285

GCT AAA GAA AGC CTC TAT TCT CAA CTG CCA ATG GAC TGT TTT ACA ATG        912
Ala Lys Glu Ser Leu Tyr Ser Gln Leu Pro Met Asp Cys Phe Thr Met
       290                 295                 300

CCA TCA TAT TCC AGA CGC ATC TCC ACA GCT ACG CCA TAT ATG AAT GGA        960
Pro Ser Tyr Ser Arg Arg Ile Ser Thr Ala Thr Pro Tyr Met Asn Gly
305                 310                 315                 320

GAA ACA TCT ACA AAA TCC CTT TGG GTT ATA AAT AGT GCA CTC AGA ATA       1008
Glu Thr Ser Thr Lys Ser Leu Trp Val Ile Asn Ser Ala Leu Arg Ile
                   325                 330                 335
```

```
AAA ATT CTT TGT GCA ACC TAT GTG AAT GTA AAT ATT CGA GAC ATT GAC      1056
Lys Ile Leu Cys Ala Thr Tyr Val Asn Val Asn Ile Arg Asp Ile Asp
            340                 345                 350

AAG ATT TAT GTT CGA ACA GGT ATC TAC CAT GGA GGA GAA CCC TTA TGT      1104
Lys Ile Tyr Val Arg Thr Gly Ile Tyr His Gly Gly Glu Pro Leu Cys
        355                 360                 365

GAT AAT GTG AAC ACT CAA AGA GTA CCT TGT TCC AAT CCC AGG TGG AAT      1152
Asp Asn Val Asn Thr Gln Arg Val Pro Cys Ser Asn Pro Arg Trp Asn
    370                 375                 380

GAA TGG CTG AAT TAC GAT ATA TAC ATT CCT GAT CTT CCT CGT GCT GCT      1200
Glu Trp Leu Asn Tyr Asp Ile Tyr Ile Pro Asp Leu Pro Arg Ala Ala
385                 390                 395                 400

CGA CTT TGC CTT TCC ATT TGT TCT GTT AAA GGC CGA AAG GGT GCT AAA      1248
Arg Leu Cys Leu Ser Ile Cys Ser Val Lys Gly Arg Lys Gly Ala Lys
                405                 410                 415

GAG GAA CAC TGT CCA TTG GCC TGG GGA AAT ATA AAC TTG TTT GAT TAC      1296
Glu Glu His Cys Pro Leu Ala Trp Gly Asn Ile Asn Leu Phe Asp Tyr
            420                 425                 430

ACA GAT ACT CTA GTA TCT GGA AAA ATG GCT TTG AAT CTT TGG CCA GTA      1344
Thr Asp Thr Leu Val Ser Gly Lys Met Ala Leu Asn Leu Trp Pro Val
        435                 440                 445

CCT CAT GGA CTA GAA GAT TTG CTG AAC CCT ATT GGT GTT ACT GGA TCA      1392
Pro His Gly Leu Glu Asp Leu Leu Asn Pro Ile Gly Val Thr Gly Ser
    450                 455                 460

AAT CCA AAT AAA GAA ACT CCA TGT TTA GAG TTG GAG TTT GAC TGG TTC      1440
Asn Pro Asn Lys Glu Thr Pro Cys Leu Glu Leu Glu Phe Asp Trp Phe
465                 470                 475                 480

AGC AGT GTG GTA AAG TTT CCA GAT ATG TCA GTG ATT GAA GAG CAT GCC      1488
Ser Ser Val Val Lys Phe Pro Asp Met Ser Val Ile Glu Glu His Ala
                485                 490                 495

AAT TGG TCT GTA TCC CGT GAA GCA GGA TTT AGT TAT TCC CAT GCA GGA      1536
Asn Trp Ser Val Ser Arg Glu Ala Gly Phe Ser Tyr Ser His Ala Gly
            500                 505                 510

CTG AGT AAC AGA CTA GCT AGA GAC AAT GAA TTA AGA GAA AAT GAT AAA      1584
Leu Ser Asn Arg Leu Ala Arg Asp Asn Glu Leu Arg Glu Asn Asp Lys
        515                 520                 525

GAA CAG CTC CGA GCA ATT TGT ACA CGA GAT CCT CTA TCT GAA ATC ACT      1632
Glu Gln Leu Arg Ala Ile Cys Thr Arg Asp Pro Leu Ser Glu Ile Thr
    530                 535                 540

GAG CAA GAG AAA GAT TTT CTG TGG AGC CAC AGA CAC TAT TGT GTA ACT      1680
Glu Gln Glu Lys Asp Phe Leu Trp Ser His Arg His Tyr Cys Val Thr
545                 550                 555                 560

ATC CCC GAA ATT CTA CCC AAA TTG CTT CTG TCT GTT AAA TGG AAC TCT      1728
Ile Pro Glu Ile Leu Pro Lys Leu Leu Leu Ser Val Lys Trp Asn Ser
                565                 570                 575

AGA GAT GAA GTA GCT CAG ATG TAC TGC TTG GTA AAA GAT TGG CCT CCA      1776
Arg Asp Glu Val Ala Gln Met Tyr Cys Leu Val Lys Asp Trp Pro Pro
            580                 585                 590

ATC AAG CCT GAA CAG GCT ATG GAG CTT CTG GAC TGC AAT TAC CCA GAT      1824
Ile Lys Pro Glu Gln Ala Met Glu Leu Leu Asp Cys Asn Tyr Pro Asp
```

595 600 605

CCT ATG GTT CGA GGT TTT GCT GTT CGG TGC TTA GAA AAA TAT TTA ACA 1872
Pro Met Val Arg Gly Phe Ala Val Arg Cys Leu Glu Lys Tyr Leu Thr
610 615 620

GAT GAC AAA CTT TCT CAG TAC CTA ATT CAG CTA GTA CAG GTA CTA AAA 1920
Asp Asp Lys Leu Ser Gln Tyr Leu Ile Gln Leu Val Gln Val Leu Lys
625 630 635 640

TAT GAA CAG TAT TTG GAT AAC CTG CTT GTG AGA TTT TTA CTC AAA AAA 1968
Tyr Glu Gln Tyr Leu Asp Asn Leu Leu Val Arg Phe Leu Leu Lys Lys
645 650 655

GCG TTA ACT AAT CAA AGG ATC GGT CAC TTT TTC TTT TGG CAT TTA AAA 2016
Ala Leu Thr Asn Gln Arg Ile Gly His Phe Phe Phe Trp His Leu Lys
660 665 670

TCT GAG ATG CAC AAT AAA ACA GTT AGT CAG AGG TTT GGC CTG CTT TTG 2064
Ser Glu Met His Asn Lys Thr Val Ser Gln Arg Phe Gly Leu Leu Leu
675 680 685

GAG TCC TAT TGC CGT GCA TGT GGG ATG TAT CTG AAG CAC CTT AAT AGG 2112
Glu Ser Tyr Cys Arg Ala Cys Gly Met Tyr Leu Lys His Leu Asn Arg
690 695 700

CAA GTT GAG GCT ATG GAA AAG CTC ATT AAC TTG ACT GAC ATT CTC AAA 2160
Gln Val Glu Ala Met Glu Lys Leu Ile Asn Leu Thr Asp Ile Leu Lys
705 710 715 720

CAA GAG AAG AAG GAT GAA ACA CAA AAG GTA CAG ATG AAG TTT TTA GTT 2208
Gln Glu Lys Lys Asp Glu Thr Gln Lys Val Gln Met Lys Phe Leu Val
725 730 735

GAG CAA ATG CGG CGA CCA GAT TTC ATG GAT GCT CTC CAG GGC TTT CTG 2256
Glu Gln Met Arg Arg Pro Asp Phe Met Asp Ala Leu Gln Gly Phe Leu
740 745 750

TCT CCT CTA AAC CCT GCT CAT CAG CTG GGA AAT CTC AGG CTT GAA GAG 2304
Ser Pro Leu Asn Pro Ala His Gln Leu Gly Asn Leu Arg Leu Glu Glu
755 760 765

TGT CGA ATT ATG TCT TCT GCA AAA AGG CCA CTG TGG TTG AAT TGG GAG 2352
Cys Arg Ile Met Ser Ser Ala Lys Arg Pro Leu Trp Leu Asn Trp Glu
770 775 780

AAC CCA GAC ATC ATG TCA GAA TTA CAC TTT CAG AAC AAT GAG ATC ATC 2400
Asn Pro Asp Ile Met Ser Glu Leu His Phe Gln Asn Asn Glu Ile Ile
785 790 795 800

TTT AAA AAT GGG GAT GAT TTA CGG CAA GAT ATG CTA ACC CTT CAG ATT 2448
Phe Lys Asn Gly Asp Asp Leu Arg Gln Asp Met Leu Thr Leu Gln Ile
805 810 815

ATT CGC ATT ATG GAA AAT ATC TGG CAA AAT CAA GGT CTT GAT CTT CGA 2496
Ile Arg Ile Met Glu Asn Ile Trp Gln Asn Gln Gly Leu Asp Leu Arg
820 825 830

ATG TTA CCT TAT GGA TGT CTG TCA ATC GGT GAC TGT GTG GGA CTT ATC 2544
Met Leu Pro Tyr Gly Cys Leu Ser Ile Gly Asp Cys Val Gly Leu Ile
835 840 845

GAG GTG GTG AGA AAT TCT CAC ACT ATA ATG CAG ATT CAG TGT AAA GGA 2592
Glu Val Val Arg Asn Ser His Thr Ile Met Gln Ile Gln Cys Lys Gly
850 855 860

```
GGC CTG AAA GGT GCA CTG CAG TTT AAC AGC CAC ACA CTC CAT CAG TGG      2640
Gly Leu Lys Gly Ala Leu Gln Phe Asn Ser His Thr Leu His Gln Trp
865                 870                 875                 880

CTC AAA GAC AAG AAC AAG GGG GAA ATA TAT GAT GCG GCC ATC GAT TTG      2688
Leu Lys Asp Lys Asn Lys Gly Glu Ile Tyr Asp Ala Ala Ile Asp Leu
                885                 890                 895

TTT ACA CGA TCA TGT GCT GGA TAT TGT GTT GCC ACC TTC ATT TTG GGA      2736
Phe Thr Arg Ser Cys Ala Gly Tyr Cys Val Ala Thr Phe Ile Leu Gly
            900                 905                 910

ATT GGA GAT CGT CAC AAT AGT AAT ATC ATG GTT AAA GAT GAT GGA CAA      2784
Ile Gly Asp Arg His Asn Ser Asn Ile Met Val Lys Asp Asp Gly Gln
        915                 920                 925

CTG TTT CAT ATA GAT TTT GGA CAC TTT TTG GAT CAC AAG AAG AAA AAA      2832
Leu Phe His Ile Asp Phe Gly His Phe Leu Asp His Lys Lys Lys Lys
    930                 935                 940

TTT GGT TAT AAA CGA GAG CGC GTG CCG TTT GTT TTG ACA CAA GAT TTC      2880
Phe Gly Tyr Lys Arg Glu Arg Val Pro Phe Val Leu Thr Gln Asp Phe
945                 950                 955                 960

TTA ATA GTG ATT AGT AAA GGA GCC CAA GAA TGC ACA AAG ACA AGA GAA      2928
Leu Ile Val Ile Ser Lys Gly Ala Gln Glu Cys Thr Lys Thr Arg Glu
                965                 970                 975

TTT GAG AGG TTT CAG GAG ATG TGT TAC AAG GCT TAT CTA GCT ATT CGG      2976
Phe Glu Arg Phe Gln Glu Met Cys Tyr Lys Ala Tyr Leu Ala Ile Arg
            980                 985                 990

CAG CAT GCC AAT CTC TTC ATA AAT CTT TTC TCA ATG ATG CTT GGC TCT      3024
Gln His Ala Asn Leu Phe Ile Asn Leu Phe Ser Met Met Leu Gly Ser
        995                 1000                1005

GGA ATG CCA GAA CTG CAA TCT TTT GAT GAT ATT GCA TAC ATT CGA AAG      3072
Gly Met Pro Glu Leu Gln Ser Phe Asp Asp Ile Ala Tyr Ile Arg Lys
    1010                1015                1020

ACC CTA GCT TTA GAT AAA ACT GAG CAA GAG GCT TTG GAG TAT TTC ATG      3120
Thr Leu Ala Leu Asp Lys Thr Glu Gln Glu Ala Leu Glu Tyr Phe Met
1025                1030                1035                1040

AAA CAA ATG AAT GAT GCA CAC CAT GGT GGC TGG ACA ACA AAA ATG GAT      3168
Lys Gln Met Asn Asp Ala His His Gly Gly Trp Thr Thr Lys Met Asp
                1045                1050                1055

TGG ATC TTC CAC ACA ATT AAG CAG CAT GCT TTG AAC TGAAATGATA           3214
Trp Ile Phe His Thr Ile Lys Gln His Ala Leu Asn
            1060                1065

ACTAAAAGCT CAGTATCTGG ATTCTACACT GCACTGTTAA TAACTGTCAA CAGGCAAAGA    3274

CTGATTGCAT AGGAATTGCA CAATCCATGA ACAGCATTAG AATTTACAGC AAGAACAGAA    3334

ATAAAATACT ATATAATTTA AATAATGTAA ACGCAAACAG GGTTTGATAG CACTAAACTA    3394

GTTCATTTCA AAATTAAGCT TTAGAATAAT GCGCAATTTC ATGTTATGCC TTAAGTCCAA    3454

AAAGGTAAAC TTTAAAGATT GTTTGTATCT TTCCTTTAAA AAAA                     3498
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1068 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Pro Pro Arg Pro Ser Ser Gly Glu Leu Trp Gly Ile His Leu Met
  1               5                  10                  15

Pro Pro Arg Ile Leu Val Glu Cys Leu Leu Pro Asn Gly Met Ile Val
             20                  25                  30

Thr Leu Glu Cys Leu Arg Glu Ala Thr Leu Ile Thr Ile Lys His Glu
         35                  40                  45

Leu Phe Lys Glu Ala Arg Lys Tyr Pro Leu His Gln Leu Leu Gln Asp
     50                  55                  60

Glu Ser Ser Tyr Ile Phe Val Ser Val Thr Gln Glu Ala Glu Arg Glu
 65                  70                  75                  80

Glu Phe Phe Asp Glu Thr Arg Arg Leu Cys Asp Leu Arg Leu Phe Gln
                 85                  90                  95

Pro Phe Leu Lys Val Ile Glu Pro Val Gly Asn Arg Glu Glu Lys Ile
            100                 105                 110

Leu Asn Arg Glu Ile Gly Phe Ala Ile Gly Met Pro Val Cys Glu Phe
        115                 120                 125

Asp Met Val Lys Asp Pro Glu Val Gln Asp Phe Arg Arg Asn Ile Leu
    130                 135                 140

Asn Val Cys Lys Glu Ala Val Asp Leu Arg Asp Leu Asn Ser Pro His
145                 150                 155                 160

Ser Arg Ala Met Tyr Val Tyr Pro Pro Asn Val Glu Ser Ser Pro Glu
                165                 170                 175

Leu Pro Lys His Ile Tyr Asn Lys Leu Asp Lys Gly Gln Ile Ile Val
            180                 185                 190

Val Ile Trp Val Ile Val Ser Pro Asn Asn Asp Lys Gln Lys Tyr Thr
        195                 200                 205

Leu Lys Ile Asn His Asp Cys Val Pro Glu Gln Val Ile Ala Glu Ala
    210                 215                 220

Ile Arg Lys Lys Thr Arg Ser Met Leu Leu Ser Ser Glu Gln Leu Lys
225                 230                 235                 240

Leu Cys Val Leu Glu Tyr Gln Gly Lys Tyr Ile Leu Lys Val Cys Gly
                245                 250                 255

Cys Asp Glu Tyr Phe Leu Glu Lys Tyr Pro Leu Ser Gln Tyr Lys Tyr
            260                 265                 270

Ile Arg Ser Cys Ile Met Leu Gly Arg Met Pro Asn Leu Met Leu Met
        275                 280                 285

Ala Lys Glu Ser Leu Tyr Ser Gln Leu Pro Met Asp Cys Phe Thr Met
    290                 295                 300

Pro Ser Tyr Ser Arg Arg Ile Ser Thr Ala Thr Pro Tyr Met Asn Gly
```

305 310 315 320

Glu Thr Ser Thr Lys Ser Leu Trp Val Ile Asn Ser Ala Leu Arg Ile
325 330 335

Lys Ile Leu Cys Ala Thr Tyr Val Asn Val Asn Ile Arg Asp Ile Asp
340 345 350

Lys Ile Tyr Val Arg Thr Gly Ile Tyr His Gly Gly Glu Pro Leu Cys
355 360 365

Asp Asn Val Asn Thr Gln Arg Val Pro Cys Ser Asn Pro Arg Trp Asn
370 375 380

Glu Trp Leu Asn Tyr Asp Ile Tyr Ile Pro Asp Leu Pro Arg Ala Ala
385 390 395 400

Arg Leu Cys Leu Ser Ile Cys Ser Val Lys Gly Arg Lys Gly Ala Lys
405 410 415

Glu Glu His Cys Pro Leu Ala Trp Gly Asn Ile Asn Leu Phe Asp Tyr
420 425 430

Thr Asp Thr Leu Val Ser Gly Lys Met Ala Leu Asn Leu Trp Pro Val
435 440 445

Pro His Gly Leu Glu Asp Leu Leu Asn Pro Ile Gly Val Thr Gly Ser
450 455 460

Asn Pro Asn Lys Glu Thr Pro Cys Leu Glu Leu Glu Phe Asp Trp Phe
465 470 475 480

Ser Ser Val Val Lys Phe Pro Asp Met Ser Val Ile Glu Glu His Ala
485 490 495

Asn Trp Ser Val Ser Arg Glu Ala Gly Phe Ser Tyr Ser His Ala Gly
500 505 510

Leu Ser Asn Arg Leu Ala Arg Asp Asn Glu Leu Arg Glu Asn Asp Lys
515 520 525

Glu Gln Leu Arg Ala Ile Cys Thr Arg Asp Pro Leu Ser Glu Ile Thr
530 535 540

Glu Gln Glu Lys Asp Phe Leu Trp Ser His Arg His Tyr Cys Val Thr
545 550 555 560

Ile Pro Glu Ile Leu Pro Lys Leu Leu Leu Ser Val Lys Trp Asn Ser
565 570 575

Arg Asp Glu Val Ala Gln Met Tyr Cys Leu Val Lys Asp Trp Pro Pro
580 585 590

Ile Lys Pro Glu Gln Ala Met Glu Leu Leu Asp Cys Asn Tyr Pro Asp
595 600 605

Pro Met Val Arg Gly Phe Ala Val Arg Cys Leu Glu Lys Tyr Leu Thr
610 615 620

Asp Asp Lys Leu Ser Gln Tyr Leu Ile Gln Leu Val Gln Val Leu Lys
625 630 635 640

Tyr Glu Gln Tyr Leu Asp Asn Leu Leu Val Arg Phe Leu Leu Lys Lys
645 650 655

Ala Leu Thr Asn Gln Arg Ile Gly His Phe Phe Phe Trp His Leu Lys

660 665 670

Ser Glu Met His Asn Lys Thr Val Ser Gln Arg Phe Gly Leu Leu Leu
675 680 685

Glu Ser Tyr Cys Arg Ala Cys Gly Met Tyr Leu Lys His Leu Asn Arg
690 695 700

Gln Val Glu Ala Met Glu Lys Leu Ile Asn Leu Thr Asp Ile Leu Lys
705 710 715 720

Gln Glu Lys Lys Asp Glu Thr Gln Lys Val Gln Met Lys Phe Leu Val
725 730 735

Glu Gln Met Arg Arg Pro Asp Phe Met Asp Ala Leu Gln Gly Phe Leu
740 745 750

Ser Pro Leu Asn Pro Ala His Gln Leu Gly Asn Leu Arg Leu Glu Glu
755 760 765

Cys Arg Ile Met Ser Ser Ala Lys Arg Pro Leu Trp Leu Asn Trp Glu
770 775 780

Asn Pro Asp Ile Met Ser Glu Leu His Phe Gln Asn Asn Glu Ile Ile
785 790 795 800

Phe Lys Asn Gly Asp Asp Leu Arg Gln Asp Met Leu Thr Leu Gln Ile
805 810 815

Ile Arg Ile Met Glu Asn Ile Trp Gln Asn Gln Gly Leu Asp Leu Arg
820 825 830

Met Leu Pro Tyr Gly Cys Leu Ser Ile Gly Asp Cys Val Gly Leu Ile
835 840 845

Glu Val Val Arg Asn Ser His Thr Ile Met Gln Ile Gln Cys Lys Gly
850 855 860

Gly Leu Lys Gly Ala Leu Gln Phe Asn Ser His Thr Leu His Gln Trp
865 870 875 880

Leu Lys Asp Lys Asn Lys Gly Glu Ile Tyr Asp Ala Ala Ile Asp Leu
885 890 895

Phe Thr Arg Ser Cys Ala Gly Tyr Cys Val Ala Thr Phe Ile Leu Gly
900 905 910

Ile Gly Asp Arg His Asn Ser Asn Ile Met Val Lys Asp Asp Gly Gln
915 920 925

Leu Phe His Ile Asp Phe Gly His Phe Leu Asp His Lys Lys Lys Lys
930 935 940

Phe Gly Tyr Lys Arg Glu Arg Val Pro Phe Val Leu Thr Gln Asp Phe
945 950 955 960

Leu Ile Val Ile Ser Lys Gly Ala Gln Glu Cys Thr Lys Thr Arg Glu
965 970 975

Phe Glu Arg Phe Gln Glu Met Cys Tyr Lys Ala Tyr Leu Ala Ile Arg
980 985 990

Gln His Ala Asn Leu Phe Ile Asn Leu Phe Ser Met Met Leu Gly Ser
995 1000 1005

Gly Met Pro Glu Leu Gln Ser Phe Asp Asp Ile Ala Tyr Ile Arg Lys

1010 1015 1020

Thr Leu Ala Leu Asp Lys Thr Glu Gln Glu Ala Leu Glu Tyr Phe Met
1025 1030 1035 1040

Lys Gln Met Asn Asp Ala His His Gly Gly Trp Thr Thr Lys Met Asp
1045 1050 1055

Trp Ile Phe His Thr Ile Lys Gln His Ala Leu Asn
1060 1065

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2199 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: SCHIZOSACCHAROMYCES POMBE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAAAATCTCA ACACATGTGA ATGATCAGAA AATTATCGCC ATAAAAGACA GAATAAGTCA 60

TCAGCGGTTG TTTCATTTCC TATATTTTTT TTTTATTTTT TTATTTTTTA ATAAGGGAAA 120

ATTTAACGTC TAAGGATACA GAAGATTGTT AGCACATTAA AGTAATAAAG GCTTAAGTAG 180

TAAGTGCCTT AGCATGTTAT TGTATTTCAA AGGACATAAT CTAAAATAAT AACAATATCA 240

TTTCTCACAA GTTATTCAAT TTTCTTTTTT TTTTCTAATA ATATCAAGAA TGTATTATTT 300

GTTTGACATA AGTCAACTAA TTTATTTAAT ATGCTGGATT AATCTTGCAG ACATGTAAAT 360

TAACAAGTTT TAGTCAAATA ACGTTGAAGT TTCAATGAAC TCAAATAATT TCTCTTTTTT 420

TTTATATAAC CATATGTCTA ATCTGATTTA TATTTTCCGC AGGATCAACT GAAGTTATGA 480

CATTTGGATT GGATCACTTA TAACCTTGGT CGCCAAATAA TACAAAAATC AGCGTTATAA 540

AACAAAGAAG GTTTTTGTTA AGAAATTAAT CCTCTTTCTT GATAAGAAAG TTGAACCGAA 600

ATTGCAGATA CTGATATATG AAAATAATAC CCACAATTTT GGGAATAGCG CAAGCCTCAA 660

TTTAAACAAT AGGTGAGGAC ACATGATAAT GACCTCAATG ATTGTTAGAA GAAAAGAGCC 720

TCATTACAAA ATCGAAAAAT GAATGGTTGG GTACAAGTTT CCAAAACATG GTAAAGTGGA 780

CTTTGCGTAT GAGACGTAAA TAGAAAAAAA CACTTGTTAT ATGTTTTCTA GAATTATTGT 840

TGTCTCTTTA TGGTTGGATG ATGCAAAATA GTAATTTCGG TTAGTTGCTG TAAAACACCA 900

CGAGACAAAT AGATATGGAT ATTTATTAAA TCAGGAAAAA CGTAACTCTC GGCTACTGGA 960

TGGTTCAGTC ACCCAACGAT TACTGGGGAG AGAAAACAGG GCAAAAGCAA AGCTTAAAGG 1020

AATCCGATTG TCATTCGGCA ATGTGCAGCG AAACTAAAAA CCGGATAATG GACCTGTTAA 1080

TCGAAACATT GAAGATATAT AAAGGAAGAG GAATCCTGGC ATATCATCAA TTGAATAAGT 1140

TGAATTAATT ATTTCAATCT CATTCTCACT TTCTGACTTA TAGTCGCTTT GTTAAATCAT 1200

AGGAATGTCT CCCTTGCCAG TACTGCTAGG GTTTTTCTTT CAAACTATGG AAGCCCATTC 1260

AAGCTGCATA TTACGATTTT GTTTTTCGCT TTTAGAAAGT GGTTTAGATG AGATAATAGA 1320

AAAATTCTTG ATCTCCGACA ACGAGTACTT TTATTTTTTT TGCTAATCAC TTTACTCAAT 1380

ATTAGCTCGA AATCGTAGAA ACGTAGACGG GTGCGGGATA CCGAGTGGTG TAGTTAAGAA 1440

TTTTTATAAA CCACGTGGCC CAAAAATATG AACCCAAAAC GTTTATACAT GAGTATACTT 1500

TAAGAAGGCT ATACCCCTTC GTGTTAGATG TAGTTTTAGC TACCCAACCC GAGTCTATGA 1560

GCTTGACTTC AGATGTAGAA GGCATTAAAT CGTTTTGAAT ATTAATTAAA AAACGATGAA 1620

AATTAAATAT TTAAAAGCAA TCATACGCTG AAAATTTAGT GCTGTGGCTA ATCCTTCAAC 1680

ATGGAAATGC CATAAAAGTG ACTTTGACAA AAAAAAAAGT ATATACAGGT AGTAAACTCA 1740

TCTACTTCAT TGACTTTGTT TACAGCATGT GGAAGGAGGA ATATTTATTG CTAAATCGTA 1800

GTTTAACATT CAATAAGTAA TACTATTGAA ATTCGACAAG ATTGGCCGCA TGGATGAAAA 1860

AGAGGCATTT TGCTTTGGGA GAATTAGTTC AAATTAGAAC TGAAAAAAAA AACTTTACGA 1920

GGCAAAAATG TCGGATTGAG ATCGTAAAAG TTCGCTCGTC GTCTTTTGCT TTGTGATTGT 1980

TTTCATGGAT ACATCTTGCT GGATATTTAA ATTTTAGTAC TATGTATAAG ATATTCTATA 2040

AATGTTTTAT CACCCAAACC TGTTAGCGCC TTCTTAATTC TATTCAATCT GGCTTTTGCT 2100

CTGAGACTAC TTCTTGGACT TTCACTACTT GTTAGTTATA CGGAATTTGT GTAATTAGAA 2160

GTGAAATAAT CCTTTCTATT AGTAATGCAA ACAAAAATC 2199

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2707 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CTCGAGCTGA AGAACCAGCG AGGCGGCGAG GCAGCCCCCG CGGCTTGCAG CGGAGGCGAC 60

AGCTCGTCTC CTGCCGTGGA GGTGTCGCCG GTGGTGGGGG GGAGAGACTT GCTCCAAAAA 120

AACGGACGTC TCCAGCTCTC CCCCCTCCCT GTTTTCCGTT AGGAATCCGG CGAGGAAATA 180

CATGCACTCG CTGAGAATCG GCGGCGCCAG GAGGCAGCGC CACAAGGTGT AGCGAGTGAG 240

TGGGGTGGGG CAAGAGGGGA CCCAGGAGTC CCCCAGGCTC CCGGCGCGCC TGCTCCTGCT 300

```
CTTCAATCCT GCCCACGGGG CGGACGGAGT GACCCCCGCC CCGACCATGG TAGTGTTCAA     360
TGGCCTTCTT AAGATCAAAA TCTGCGAGGC GGTGAGCTTG AAGCCCACAG CCTGGTCGCT     420
GCGCCATGCG GTGGGACCCC GGCCACAGAC GTTCCTTTTG GACCCCTACA TTGCCCTTAA     480
CGTGGACGAC TCGCGCATCG GCCAAACAGC CACCAAGCAA AAGACCAACA GCCCGGCCTG     540
GCACGATGAG TTCGTCACCG ATGTGTGCAA TGGGCGCAAG ATCGAGCTGG CTGTCTTTCA     600
CGACGCTCCT ATCGGCTACG ACGACTTCGT GGCCAACTGC ACCATCCAGT TCGAGGAGCT     660
GCTGCAGAAT GGGAGCCGTC ACTTCGAGGA CTGGATTGAC CTGGAGCCAG AAGGAAAAGT     720
GTACGTGATC ATCGATCTCT CGGGATCATC GGGTGAAGCC CCTAAAGACA ATGAAGAACG     780
AGTGTTCAGG GAGCGTATGC GGCCAAGGAA GCGGCAAGGG GCTGTCAGGC GCAGGGTCCA     840
CCAGGTCAAT GGCCACAAGT TCATGGCCAC CTACTTGCGG CAACCCACCT ACTGCTCCCA     900
CTGCAGAGAT TTCATCTGGG GTGTCATAGG AAAACAGGGA TATCAATGTC AAGTTTGCAC     960
TTGCGTTGTC CACAAGCGAT GTCATGAGCT CATTATTACA AAGTGCGCTG GGCTGAAGAA    1020
ACAGGAAACC CCTGACGAGG TGGGCTCCCA ACGGTTCAGC GTCAACATGC CCCACAAGTT    1080
CGGGATCCAC AACTACAAGG TCCCCACGTT CTGTGACCAC TGTGGGTCCC TGCTCTGGGG    1140
CCTCTTGCGG CAGGGCTTGC AGTGTAAAGT CTGCAAAATG AATGTTCACC GGCGATGTGA    1200
GACCAACGTG GCTCCCAACT GTGGGGTAGA CGCCAGAGGA ATTGCCAAAG TGCTGGCTGA    1260
CCTCGGTGTT ACTCCAGACA AAATCACCAA CAGTGGCCAA AGGAGGAAAA AGCTCGCTGC    1320
TGGTGCTGAG TCCCCACAGC CGGCTTCTGG AAACTCCCCA TCTGAAGACG ACCGATCCAA    1380
GTCAGCGCCC ACCTCCCCTT GTGACCAGGA ACTAAAAGAA CTTGAAAACA ACATCCGGAA    1440
GGCCTTGTCA TTTGACAACC GAGGAGAGGA GCACCGAGCG TCGTCGGCCA CCGATGGCCA    1500
GCTGGCAAGC CCCGGAGAGA ATGGGGAAGT CCGGCCAGGC CAGGCCAAGC GCTTGGGGCT    1560
GGATGAGTTC AACTTCATCA AAGTGTTGGG CAAAGGCAGC TTTGGCAAGG TCATGTTGGC    1620
GGAACTCAAA GGCAAAGATG AAGTCTACGC TGTGAAGGTC TTGAAGAAGG ACGTTATCCT    1680
ACAAGACGAT GATGTGGACT GCACAATGAC AGAGAAGAGG ATTTTGGCTC TGGCTCGGAA    1740
ACACCCTTAT CTAACCCAAC TCTATTGCTG CTTCCAGACC AAGGACCGCC TCTTCTTCGT    1800
CATGGAATAT GTAAATGGTG GAGACCTCAT GTTCCAGATT CAGCGGTCCC GAAAATTTGA    1860
TGAGCCTCGT TCTCGGTTCT ATGCCGCAGA GGTCACATCG GCCCTCATGT TTCTCCACCA    1920
GCATGGAGTG ATCTACAGGG ATTTGAAACT GGACAACATC CTTCTAGATG CAGAAGGCCA    1980
CTGCAAGCTG GCTGACTTTG GGATGTGCAA GGAAGGGATT ATGAATGGTG TGACAACTAC    2040
CACCTTCTGT GGGACTCCTG ACTACATAGC TCCAGAGATC CTACAGGAGT TGGAGTACGG    2100
CCCCTCAGTG GACTGGTGGG CCCTGGGGGT GCTGATGTAC GAGATGATGG CTGGGCAGCC    2160
CCCCTTTGAA GCTGACAACG AGGACGACTT GTTCGAATCC ATCCTTCATG ATGATGTTCT    2220
CTATCCTGTC TGGCTTAGCA AGGAAGCTGT CAGCATCCTG AAAGCTTTCA TGACCAAGAA    2280
```

CCCGCACAAG CGCCTGGGCT GTGTGGCAGC GCAGAACGGG GAGGACGCCA TCAAGCAACA 2340

TCCATTCTTC AAGGAGATTG ACTGGGTACT GCTGGAGCAG AAGAAAATCA AGCCCCCCTT 2400

CAAGCCGAGA ATTAAAACCA AAAGAGATGT CAATAACTTT GACCAAGACT TTACGCGGGA 2460

AGAGCCAATA CTTACACTTG TGGATGAAGC AATCATTAAG CAGATCAACC AGGAAGAATT 2520

CAAAGGCTTC TCCTACTTTG GTGAAGACCT GATGCCCTGA GAGGCTGCTT CGGATGGAGG 2580

GAGCTCATGC TGCAAGGACG GTGTTGAGAT ACTCCCAAGC TGCAGAGGCT CCGAAGGTCT 2640

CAACTCCTCC TCCTCCTCCC CCTCCCCAGA GCCCCAGTCC CATGTCCACT CTCTTATTTA 2700

TTGCATT 2707

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2167 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGCCCCTGTT CTGCAGAAAG GGGGCTCTGA GGCAGAAGGT GGTCCATGAG GTCAAGAGCC 60

ACAAGTTCAC CGCTCGCTTC TTCAAGCAGC CGACCTTCTG CAGCCACTGC ACTGACTTCA 120

TATGGGGGAT TGGAAAACAG GGTCTGCAAT GTCAAGTCTG CAGTTTTGTG GTTCATCGAC 180

GATGCCACGA GTTTGTGACC TTCGAGTGTC CAGGCGCTGG GAAGGGCCCC CAGACGGACG 240

ATCCCCGGAA CAAGCACAAG TTCCGTCTGC ACAGCTACAG CAGCCCCACC TTCTGCGACC 300

ACTGTGGCTC CCTGCTCTAC GGGCTGGTGC ACCAGGGCAT GAAGTGTTCT TGCTGCGAGA 360

TGAACGTGCA CCGGCGCTGT GTGCGCAGCG TGCCCTCTCT GTGCGGCGTG GACCACACGG 420

AGCGCCGGGG CCGCCTGCAG CTGGAGATCC GGGCGCCCAC TTCCGATGAG ATCCACGTTA 480

CGGTTGGCGA GGCCCGGAAC CTCATCCCAA TGGACCCCAA CGGTCTCTCC GATCCCTATG 540

TGAAGCTGAA GCTCATCCCA GACCCTCGGA ATTTGACCAA GCAGAAGACC CGCACGGTGA 600

AAGCTACGCT AAACCCTGTG TGGAACGAGA CCTTTGTGTT CAACCTGAAG CCGGGGGACG 660

TGGAGCGCCG GCTCAGCGTG GAGGTGTGGG ACTGGGACCG GACCTCCCGA AACGACTTCA 720

TGGGCGCCAT GTCCTTCGGC GTCTCGGAGC TGCTCAAGGC GCCGGTGGAC GGCTGGTACA 780

AGTTACTGAA CCAGGAGGAG GGCGAGTATT ACAATGTGCC GGTGGCTGAC GCCGACAACT 840

GCAACCTCCT CCAGAAGTTC GAGGCCTGTA ACTACCCCCT GGAACTATAC GAGAGGGTGC 900

GGACGGGTCC CTCTTCATCT CCCATCCCCT CCCCATCCCC CAGTCCCACC GACTCCAAGC 960

GCTGTTTCTT CGGGGCCAGC CCTGGACGAC TGCACATCTC CGACTTCAGC TTCCTCATGG 1020

TTCTAGGAAA AGGCAGTTTT GGGAAGGTGA TGCTGGCCGA GCGCCGGGGC TCCGATGAGC 1080

TCTACGCCAT CAAGATCCTG AAGAAAGACG TGATCGTCCA GGATGACGAC GTGGACTGCA 1140

CCCTGGTGGA GAAACGCGTG CTGGCTCTGG GGGGCCGAGG CCCGGGAGGC CGGCCGCACT 1200

TCCTCACCCA GCTTCACTCC ACCTTCCAGA CCCCGGATCG CCTGTATTTT GTGATGGAGT 1260

ATGTCACCGG GGGCGACTTG ATGTACCACA TTCAACAGCT GGGCAAGTTT AAGGAACCCC 1320

ACGCAGCGTT CTACGCTGCA GAAATCGCCA TCGGCCTCTT CTTCCTCCAT AACCAGGGCA 1380

TTATCTATCG GGACCTGAAA CTGGACAACG TGATGCTGGA TGCCGAAGGA CACATCAAAA 1440

TCACCGACTT CGGCATGTGT AAGGAGAACG TCTTTCCCGG GAGTACCACT CGCACCTTCT 1500

GCGGGACCCC GGACTACATA GCCCCCGAGA TCATTGCCTA CCAACCCTAT GGGAAGTCTG 1560

TGGATTGGTG GTCCTTTGGG GTTCTGCTCT ACGAGATGTT GGCAGGACAG CCCCCCTTTG 1620

ATGGAGAAGA TGAGGAGGAG CTGTTTCAAG CCATCATGGA ACAAACTGTC ACCTACCCCA 1680

AGTCGCTTTC CCGGGAAGCT GTGGCCATCT GCAAGGGGTT CCTCACCAAG CACCCGGCCA 1740

AGCGCCTGGG CTCAGGCCCC GATGGAGAGC CCACCATCCG CGCTCACGGC TTTTTCCGCT 1800

GGATCGACTG GGACAGGCTG GAACGATTAG AGATCGCGCC TCCGTTCAGA CCCCGCCCGT 1860

GTGGCCGCAG CGGCGAGAAC TTCGACAAGT TCTTCACTCG GGCGGCGCCG GCGCTGACAC 1920

CCCCTGACCG CCTGGTTCTG GCCAGCATCG ACCAGGCTGA GTTCCAGGGC TTCACCTATG 1980

TCAACCCGGA TTTCGTGCAC CCGGATGCCC GCAGCCCCAT CAGCCCAACG CCTGTGCCAG 2040

TCATGTAATC CCACCTGCCG CCACCAGGCG TCCCCACGGC TCCCTCCTCC GCCCCGGCTT 2100

TGGCCCTCGC CTCACCATGC CACCCGCCTT TCCAATTCTA GATATGGCTC CCCAGCGTTC 2160

TGGCCTC 2167

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2891 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GGCGGCGGCC GCGGGGATCC CGCGAGCGGC CCCTGAACAT CTACCCTTCT TGCCGGGACC 60

CGGGAGGTCC CCACTGGCCT CCGGGCCCGT CCTGATCAGA CTCGTGTCGA CCTCCCCGTC 120

CACGCGCATC CGGGAGAGCC GCGCCACGAG ACGGACCCGG GCCCGCCGGG ACCCCTGGTG 180

TCTGGCCCTG CGTCGAGAGG CTGGTGACTG CCACCCATAA GCTCCAGCTT CAGCCTCGGC 240

TTACTCCCCT CAGGGGCTTG CAGGCTGAGG CCTGCCCTCG GACGCGGCTG ACCAGCCTCT 300

CCCTCTCTTC CACACTTTGG ACTTCTCTTT GGACCTCCTA AAAAGGCTCC ATCATGGCAC 360

CGTTCCTGCG CATCTCCTTC AATTCCTATG AGCTGGGCTC CCTGCAGGCG GAGGACGACG 420

CAAGCCAGCC TTTCTGTGCC GTGAAGATGA AGGAGGCACT CACCACAGAC CGAGGGAAGA 480

CTCTGGTACA GAAGAAGCCC ACAATGTACC CTGAGTGGAA GTCAACATTC GACGCCCACA 540

TCTATGAAGG CCGTGTCATC CAGATCGTGC TGATGCGGGC AGCTGAAGAC CCCATGTCGG 600

AGGTGACCGT GGGCGTGTCA GTGCTGGCTG AGCGCTGCAA GAAGAACAAC GGCAAGGCTG 660

AGTTCTGGCT GGACCTGCAG CCTCAGGCCA AGGTGCTGAT GTGTGTGCAG TATTTCCTGG 720

AGGATGGGGA TTGCAAACAG TCCATGCGTA GTGAGGAGGA GGCCATGTTC CCAACTATGA 780

ACCGCCGTGG AGCCATTAAA CAGGCCAAGA TTCACTACAT CAAGAACCAC GAGTTCATCG 840

CCACCTTCTT TGGGCAGCCC ACCTTCTGTT CTGTGTGCAA AGAGTTTGTC TGGGGCCTCA 900

ACAAGCAAGG CTACAAATGC AGGCAATGCA ACGCTGCCAT CCATAAGAAA TGCATCGACA 960

AGATTATCGG CCGCTGCACT GGCACTGCTA CCAATAGCCG GGACACCATC TTCCAGAAAG 1020

AACGCTTCAA CATCGACATG CCTCACCGAT TCAAGGTCTA TAACTACATG AGCCCCACCT 1080

TCTGTGACCA CTGTGGCACT TTGCTCTGGG GATTGGTGAA ACAGGGATTA AAGTGTGAAG 1140

ACTGCGGCAT GAATGTGCAC CACAAATGCC GGGAGAAGGT GGCCAACCTG TGTGGTATCA 1200

ACCAAAAGCT CTTGGCTGAG GCCTTGAACC AAGTGACCCA GAAAGCTTCC CGGAAGCCAG 1260

AGACACCAGA GACTGTCGGA ATATACCAGG GATTCGAGAA GAAGACAGCT GTCTCTGGGA 1320

ATGACATCCC AGACAACAAC GGGACCTATG GCAAGATCTG GGAGGGGAGC AACCGGTGCC 1380

GCCTTGAGAA CTTCACCTTC CAGAAAGTAC TTGGCAAAGG CAGCTTTGGC AAGGTACTGC 1440

TTGCAGAACT GAAGGGCAAG GAAAGGTACT TTGCAATCAA GTACCTGAAG AAGGACGTGG 1500

TGTTGATCGA CGATGACGTG GAGTGCACCA TGGTGGAGAA GCGGGTGCTG GCGCTCGCCT 1560

GGGAGAATCC CTTCCTCACC CATCTCATCT GTACCTTCCA GACCAAGGAC CACCTCTTCT 1620

TTGTGATGGA GTTCCTCAAT GGGGGCGATC TGATGTTCCA CATTCAGGAC AAAGGCCGCT 1680

TCGAACTCTA CCGGGCTACG TTTTATGCAG CTGAGATCAT CTGCGGACTG CAGTTTCTAC 1740

ATGGCAAAGG CATCATTTAC AGGGACCTCA AGCTAGACAA TGTAATGCTG GACAAGGATG 1800

GCCACATCAA GATTGCTGAC TTCGGGATGT GCAAAGAGAA TATATTTGGG GAGAACCGGG 1860

CCAGCACATT CTGCGGCACT CCTGACTACA TCGCCCCTGA GATCCTGCAG GGCCTGAAGT 1920

ACTCATTTTC CGTGGACTGG TGGTCTTTTG GGGTCCTCCT CTATGAGATG CTCATTGGCC 1980

AGTCCCCCTT CCATGGTGAT GATGAGGACG AGCTCTTTGA GTCCATCCGG GTGGACACAC 2040

CACACTACCC GCGCTGGATC ACCAAGGAGT CCAAGGACAT CATGGAGAAG CTCTTCGAGA 2100

GGGACCCTGC CAAGAGGCTG GGAGTAACAG CAAACATCAG GCTTCACCCC TTTTTCAAGA 2160

CTATCAACTG GAACCTGCTG GAGAAGCGGA AGGTGGAGCC CCCCTTTAAG CCCAAAGTGA 2220

AATCCCCTTC AGACTACAGC AACTTTGACC CAGAGTTCCT GAATGAGAAA CCCCAACTTT 2280

CCTTCAGTGA CAAGAACCTC ATCGACTCTA TGGACCAGAC AGCCTTCAAG GGCTTCTCCT 2340

TTGTGAACCC CAAATATGAG CAATTCCTGG AATAGTGAGC TCCCAGACCT GCTTTTAATG 2400

CCCCGGCAGA GTAGGCCCAT CTGCCCTGGT TTGCATCCTC ACTGCCCATG AAGAAGAGTG 2460

GGTGACTGGT GATTCCTGCT GCTGCCCCCT CTTCCTCGGA GAGTCTGGCT CCTGTTGGCT 2520

GGGCTCACAG TACTTCCTCT GTGAACTGTT TGTGAATTTG CCTTCCTTTT GCCATCGGAG 2580

GGAAACTGTA AATCCTGTGT GTCATTACTT GAATGTAGTT ATTGAAATAT ATATTATATA 2640

TATGCACATA TATATAATAG GCTGTATATA TTGCTCAGTA TAGAAAGCAT GTAGGAGACT 2700

GGTGATGTGT TGACCTTTTT TTAAAAAAAA CCATATGTAT ACGTGTGTAT GTATACATCT 2760

ACACACGTAT ACATATATGT ATGTATGTAT GTATGTATGT ATGTATATAT GACCAAAAGA 2820

AAAGAGAGCA CAAGCTACCT GAACCACAGG ATTGTTTATG TGTGTATAAA TAAACACTGA 2880

ATGGTAAAAA A 2891

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2176 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TCCGGGTTCC CCAGTGCCAG CCAGCGCGGC CCCCTCGGGG CTCCGGCAGC AGCGCCGGCA 60

TGTCGTCCGG CACGATGAAG TTCAATGGCT ATCTGAGGGT CCGCATCGGA GAGGCTGTAG 120

GGCTGCAGCC CACCCGCTGG TCCCTGCGGC ACTCGCTCTT CAAAAGGGC CACCAGCTGC 180

TGGACCCCTA CCTGACGGTG AGCGTAGACC AGGTACGCGT GGGCCAGACC AGCACAAAGC 240

AGAAGACCAA CAAACCCACC TACAACGAGG AGTTCTGCGC CAATGTCACC GACGGCGGCC 300

ACCTGGAGCT AGCCGTCTTC CACGAGACGC CCCTGGGTTA TGACCACTTT GTGGCCAACT 360

GCACGCTGCA GTTCCAGGAG CTGTTGCGCA CGGCTGGTAC CTCGGACACC TTCGAGGGCT 420

GGGTGGATCT GGAGCCTGAG GGGAAAGTGT TTGTGGTAAT AACCCTAACA GGGAGTTTCA 480

CTGAAGCCAC TCTCCAGAGA GACCGCATCT TCAAGCATTT TACCAGGAAG CGCCAAAGGG 540

CTATGCGAAG ACGAGTCCAT CAAGTGAACG GACATAAGTT CATGGCCACG TACCTGAGGC 600

AGCCCACCTA CTGCTCTCAT TGCCGAGAGT TCATCTGGGG AGTATTTGGG AAACAGGGTT 660

ATCAATGCCA AGTGTGCACC TGCGTCGTCC ATAAACGCTG CCATCATCTA ATTGTTACAG 720

CCTGCACTTG CCAAAACAAT ATTAACAAAG TGGATGCCAA GATTGCAGAA CAGCGGTTTG 780

GCATCAACAT CCCACACAAG TTCAACGTTC ACAATTACAA GGTGCCCACG TTCTGTGACC 840

ACTGTGGCTC CCTGCTCTGG GGGATAATGC GACAAGGACT TCAGTGTAAA ATATGTAAGA 900

TGAATGTACA TATTCGGTGT CAGGCGAACG TGGCCCCAAA CTGCGGGGTG AATGCCGTGG 960

AGCTTGCCAA GACCCTGGCA GGGATGGGTC TCCAACCCGG AAATATTTCT CCAACCTCGA 1020

AACTCATTTC CAGGTCGACA CTAAGACGGC AGGGAAAGGA GGGCTCCAAA GAAGGAAATG 1080

GGATCGGTGT TAACTCTTCC AGCAGATTCG GCATCGACAA CTTTGAGTTC ATCCGGGTGT 1140

TGGGGAAGGG GAGCTTCGGG AAGGTGATGC TTGCCAGGAT AAAGGAGACA GGAGAACTGT 1200

ACGCCGTGAA GGTGCTGAAG AAGGACGTGA TTCTGCAGGA TGATGATGTA GAGTGCACCA 1260

TGACTGAGAA GAGGATCCTG TCCTTGGCTC GCAACCACCC CTTCCTCACC CAGCTCTTCT 1320

GCTGCTTTCA GACTCCAGAC CGTCTGTTCT TTGTCATGGA GTTTGTGAAC GGAGGCGACC 1380

TGATGTTCCA CATCCAAAAG TCCCGTCGTT TCGATGAAGC CCGTGCTCGT TTCTACGCCG 1440

CGGAGATCAT TTCTGCACTC ATGTTCCTAC ATGAGAAAGG TATCATCTAT AGAGACTTGA 1500

AACTGGACAA TGTGCTATTG GACCACGAAG GTCACTGTAA ACTGGCCGAT TTTGGAATGT 1560

GCAAGGAGGG GATTTGTAAT GGGGTCACCA CAGCCACCTT CTGCGGTACA CCTGACTACA 1620

TTGCCCCAGA GATCCTTCAG GAGATGCTGT ATGGACCTGC AGTAGACTGG TGGGCCATGG 1680

GCGTGTTGCT TTATGAGATG CTGTGCGGAC ATGCGCCCTT TGAGGCTGAA AATGAAGATG 1740

ACCTTTTTGA GGCCATACTG AATGATGAAG TCGTCTACCC CACCTGGCTC CATGAAGATG 1800

CCAGAGGGAT CCTCAAGTCT TTCATGACCA AGAACCCCAC CATGCGCTTG GGCAGCCTGA 1860

CTCAGGGAGG AGAGCATGAG ATCCTGAGAC ACCCTTTCTT TAAGGAAATC GACTGGGCCC 1920

AGTTGAACCA TCGCCAGTTA GAGCCGCCTT TCCGACCTAG AATCAAATCC CGAGAAGATG 1980

TCAGCAATTT TGACCCAGAC TTTATAAAAG AAGAGCCCGT CTTAACTCCG ATTGATGAGG 2040

GACATCTTCC TATGATTAAC CAGGATGAGT TTAGAAACTT TTCCTATGTG TCACCGGAAT 2100

TGCAACTGTA GCCTTATGGG GAGTCAGAAC CAAAGGGGAA GGTGGATTTC TCCAGGAATT 2160

TCTTATGTGG GAATTC 2176

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met Asp Trp Ile Phe His Thr

1 5

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

AARATGGAYT GGATHTTYCA YAC 23

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Asp Asp Gly Gln Leu Phe His Ile Asp Phe Gly His Phe
1 5 10

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GATGATGGCC ARCTGTTYCA YATWGAYTTT GGCCAYTT 38

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

AATTCACACA CTGGCATGCC GAT 23

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GACTCGAGTC GACATCGATT TTTTTTTTTT TTTTT 35

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TTTAAGCTTA GGCATTCTAA AGTCACTATC ATCCC 35

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GACTCGAGTC GACATCGA 18

CLAIMS

1. A eukaryotic cell transformed with a DNA construct comprising a coding sequence encoding a polypeptide having the activity of a mammalian phospholipid kinase or a mammalian protein kinase activated by a phospholipid or its metabolite and regulatory elements to allow transcription of the coding sequence in the said cell wherein the regulatory elements include a repressible or inducible promoter and the expression of the said coding sequence is lethal or growth inhibitory to the cell.

2. A cell according to Claim 1 wherein the cell is a yeast cell.

3. A cell according to Claim 2 wherein the yeast is *Schizosaccharomyces*.

4. A cell according to Claim 3 wherein the promoter is the *nmt* promoter.

5. A *Schizosaccharomyces* cell transformed with a DNA construct comprising a coding sequence encoding a polypeptide having the activity of a mammalian phospholipid kinase or a mammalian protein kinase activated by a phospholipid or its metabolite and regulatory elements to allow transcription of the coding sequence in the said cell wherein the regulatory elements include a constitutive promoter and the expression of the said coding sequence is lethal or growth inhibitory to the cell.

6. A *Schizosaccharomyces* cell according to Claim 5 wherein the promoter is the *adh* promoter.

7. A cell according to any one of the preceding claims wherein the

phospholipid kinase is an inositol phospholipid kinase.

8. A cell according to any one of Claims 1 to 6 wherein the protein kinase activated by a phospholipid or its metabolite is a protein kinase C.

9. A cell according to Claim 7 wherein the phospholipid kinase is selected from the group consisting of phosphatidyl inositol 3-kinase, phosphatidyl inositol 4-kinase and phosphatidyl inositol-5-kinase.

10. A cell according to Claim 9 wherein the phospholipid kinase is phosphatidyl inositol 3-kinase.

11. A cell according to Claim 8 wherein the protein kinase C is selected from any one of PKC-γ, PKC-δ, PKC-η or PKC-ϵ.

12. An assay for detecting whether a compound is involved in cell growth regulation, the assay comprising (1) a cell according to any one of the preceding claims, (2) a container for the said cell, (3) a growth medium for the said cell and (4) means to detect the viability of the cell.

13. A kit comprising a eukaryotic cell as defined in Claim 1 and culture medium such that the cell will divide and grow.

14. A method for assaying for a compound that is involved in cell growth regulation the method comprising (1) culturing a cell as defined in Claim 1, (2) adding a compound and (3) determining the cell growth rate in the presence of the compound.

15. A compound identified by the assay of Claim 12 or the method of Claim 14 as being involved in cell growth regulation.

SUBSTITUTE SHEET

```
M P P R P S S G E L W G I H L M P P R I L V E C L L P N G M I V T L E C L R E A   40
ATGCCTCCAAGACCATCATCAGGTGAACTGTGGGGCATCCACTTGATGCCCCCAAGAATCCTAGTAGAATGTTTACTACCAAATGGGATGATAGTGACTTTAGAATGCCTCCGTGAGGCT  120

T L I T I K H E L F K E A R K Y P L H Q L L Q D E S S Y I F V S V T Q E A E R E   80
ACGTTAATAACGATAAAGCATGAACTATTTAAAGAAGCAAGAAAATACCCTCTCCATCAACTTCTTCAAGATGAATCTTCTTACATTTTCGTAAGTGTTACCCAAGAAGCAGAAAGGGAA  240

E F F D E T R R L C D L R L F Q P F L K V I E P V G N R E E K I L N R E I G F A  120
GAATTTTTTGATGAAACAAGACGACTTTGTGACCTTCGGCTTTTTCAACCCTTTTTAAAAGTAATTGAACCAGTAGGCAACCGTGAAGAAAAGATCCTCAATCGAGAAATTGGTTTTGCT  360

I G H P V C E F D H V K D P E V Q D F R R N I L N V C K E A V D L R D L N S P H  160
ATCGGCATGCCAGTGTGTGAATTCGATATGGTTAAAGATCCAGAAGTACAGGACTTCCGAAGAAATATTCTCAATGTTTGTAAAGAAGCTGTGGATCTTAGGGATCTTAATTCACCTCAT  480
                A
S R A H Y V Y P P N V E S S P E L P K H I Y N K L D K G Q I I V V I W V I V S P  200
AGTAGAGCAATGTATGTTTATCCTCCAAATGTAGAATCTTCACCAGAACTGCCAAAGCACATATATAATAAATTGGATAAAGGGCAAATAATAGTGGTGATTTGGGTAATAGTTTCTCCA  600

N N D K Q K Y T L K I N H D C V P E Q V I A E A I R K K T R S M L L S S E Q L K  240
AATAATGACAAACAGAAGTATACTCTGAAAATCAACCATGACTGTGTGCCAGAACAAGTAATTGCTGAAGCAATCAGGAAAAAAACTCGAAGTATGTTGCTATCATCTGAACAACTAAAA  720

L C V L E Y Q G K Y I L K V C G C D E Y F L E K Y P L S Q Y K Y I R S C I M L G  280
CTCTGTGTTTTAGAATATCAGGGCAAGTATATTTTAAAAGTGTGTGGATGTGATGAATACTTCCTAGAAAAATATCCTCTGAGTCAGTATAAGTATATAAGAAGCTGTATAATGCTTGGG  840

R M P N L M L M A K E S L Y S Q L P M D C F T M P S Y S R R I S T A T P Y M N G  320
AGGATGCCCAATTTGATGCTGATGGCTAAAGAAAGCCTCTATTCTCAACTGCCAATGGACTGTTTTACAATGCCATCATATTCCAGACGCATCTCCACAGCTACGCCATATATGAATGGA  960
                B
E T S T K S L W V I N S A L R I K I L C A T Y V N V N I R D I D K I Y V R T G I  360
GAAACATCTACAAAATCCCTTTGGGTTATAAATAGTGCACTCAGAATAAAAATTCTTTGTGCAACCTATGTGAATGTAAATATTCGAGACATTGACAAGATTTATGTTCGAACAGGTATC 1080

Y M G G E P L C D N V N T Q R V P C S N P R W N E W L N Y D I Y I P D L P R A A  400
TACCATGGAGGAGAACCCTTATGTGATAATGTGAACACTCAAAGAGTACCTTGTTCCAATCCCAGGTGGAATGAATGGCTGAATTACGATATATACATTCCTGATCTTCCTCGTGCTGCT 1200
```

Fig. 1

```
R L C L S I C S V K G R K G A K E E H C P L A W G N I N L F D Y T D T L V S G K   440
CGACTTTGCCTTTCCATTTGTTCTGTTAAAGGCCGAAAGGGTGCTAAAGAGGAACACTGTCCATTGGCCTGGGGAAATATAAACTTGTTTGATTACACAGATACTCTAGTATCTGGAAAA  1320
                                C
M A L N L W P V P H G L E D L L N P I G V T G S N P N K E T P C L E L E F D W F   480
ATGGCTTTGAATCTTTGGCCAGTACCTCATGGACTAGAAGATTTGCTGAACCCTATTGGTGTTACTGGATCAAATCCAAATAAAGAAACTCCATGTTTAGAGTTGGAGTTTGACTGGTTC  1440

S S V V K F P D M S V I E E H A N W S V S R E A G F S Y S H A G L S N R L A R D   520
AGCAGTGTGGTAAAGTTTCCAGATATGTCAGTGATTGAAGAGCATGCCAATTGGTCTGTATCCCGTGAAGCAGGATTTAGTTATTCCCATGCAGGACTGAGTAACAGACTAGCTAGAGAC  1560

N E L R E N D K E Q L R A I C T R D P L S E I T E Q E K D F L W S H R H Y C V T   560
AATGAATTAAGAGAAAATGATAAAGAACAGCTCCGAGCAATTTGTACACGAGATCCTCTATCTGAAATCACTGAGCAAGAGAAAGATTTTCTGTGGAGCCACAGACACTATTGTGTAACT  1680

I P E I L P K L L L S V K W N S R D E V A Q M Y C L V K D W P P I K P E Q A M E   600
ATCCCCGAAATTCTACCCAAATTGCTTCTGTCTGTTAAATGGAACTCTAGAGATGAAGTAGCTCAGATGTACTGCTTGGTAAAAGATTGGCCTCCAATCAAGCCTGAACAGGCTATGGAG  1800
                                                D
L L D C N Y P D P M V R G F A V R C L E K Y L T D D K L S Q Y L I Q L V Q V L K   640
CTTCTGGACTGCAATTACCCAGATCCTATGGTTCGAGGTTTTGCTGTTCGGTGCTTAGAAAAATATTTAACAGATGACAAACTTTCTCAGTACCTAATTCAGCTAGTACAGGTACTAAAA  1920
                                E
Y E Q Y L D N W L V R F L L K K A L T N Q R I G H F F F W H L K S E M H N K T V   680
TATGAACAGTATTTGGATAACCTGCTTGTGAGATTTTTACTCAAAAAAGCCGTTAACTAATCAAAGGATCGGTCACTTTTTCTTTTGGCATTTAAAATCTGAGATGCACAATAAAACAGTT  2040
        F                                                               G
S Q R F C L L L E S Y C R A C G M Y L K H L N R Q V E A H E K L I N L T D I L K   720
AGTCAGAGGTTTGGCCTGCTTTTGGAGTCCTATTGCCGTGCATGTGGGATGTATCTGAAGCACCTTAATAGGCAAGTTGAGGCTATGGAAAAGCTCATTAACTTGACTGACATTCTCAAA  2160

Q E K K D E T Q K V Q M K F L V E Q M R R P D F M D A L Q G F L S P L N P A H Q   760
CAAGAGAAGAAGGATGAAACACAAAAGGTACAGATGAAGTTTTTAGTTGAGCAAATGCGGCGACCAGATTTCATGGATGCTCTCCAGGGCTTTCTGTCTCCTCTAAACCCTGCTCATCAG  2280
```

Fig.1

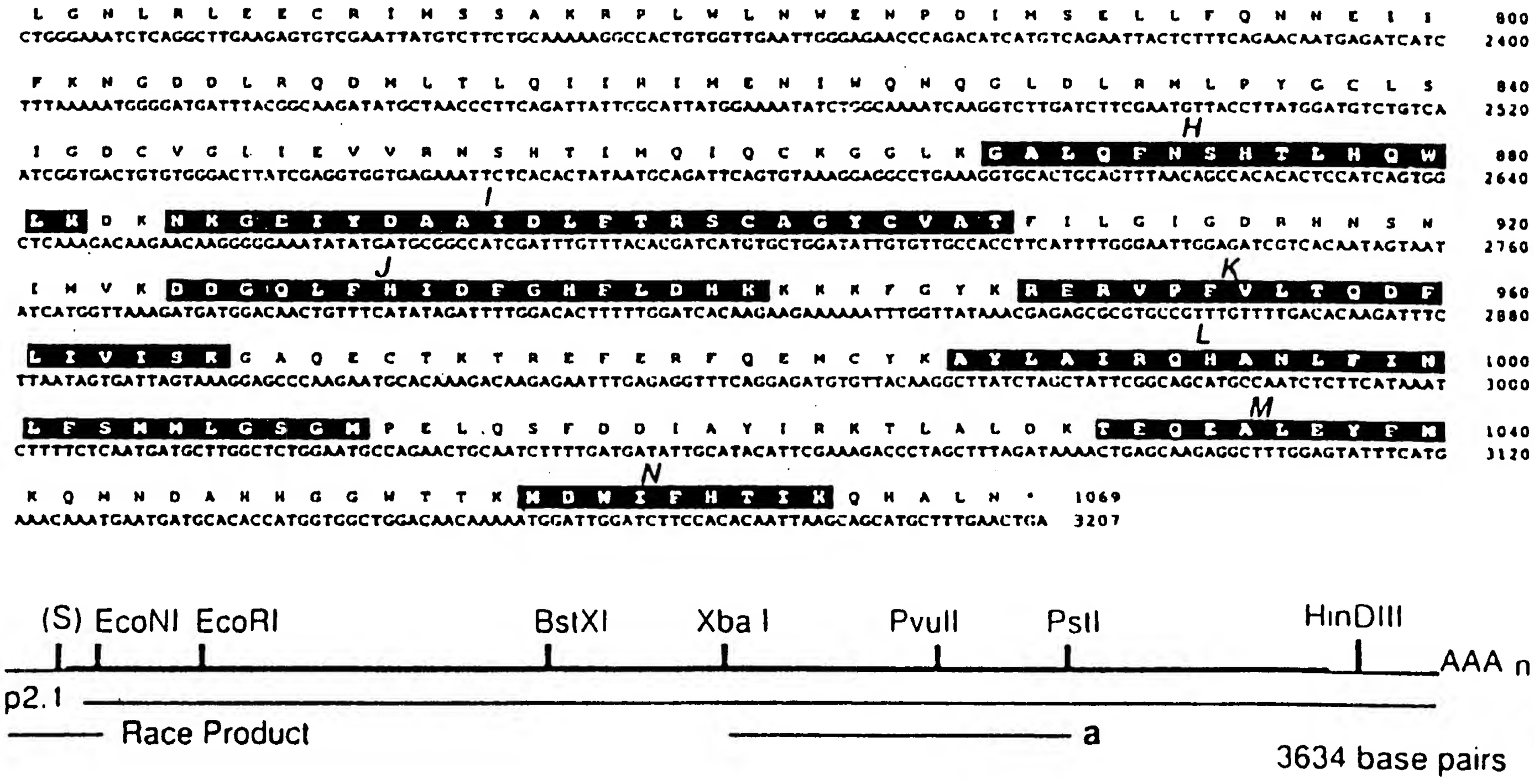

L G H L R L E E C R I M S S A K R P L W L N W E N P D I M S E L L F Q N N E I I 800
CTGGGAAATCTCAGGCTTGAAGAGTGTCGAATTATGTCTTCTGCAAAAAGGCCACTGTGGTTGAATTGGGAGAACCCAGACATCATGTCAGAATTACTCTTTCAGAACAATGAGATCATC 2400
F K N G D D L R Q D M L T L Q I I H I M E N I W Q N Q G L D L R M L P Y G C L S 840
TTTAAAAATGGGGATGATTTACGGCAAGATATGCTAACCCTTCAGATTATTCGCATTATGGAAAATATCTGGCAAAATCAAGGTCTTGATCTTCGAATGTTACCTTATGGATGTCTGTCA 2520
H
I G D C V G L I E V V R N S H T I M Q I Q C K G G L K G A L Q F N S H T L H Q W 880
ATCGGTGACTGTGTGGGACTTATCGAGGTGGTGAGAAATTCTCACACTATAATGCAGATTCAGTGTAAAGGAGGCCTGAAAGGTGCACTGCAGTTTAACAGCCACACACTCCATCAGTGG 2640
I
L K D K N K G E I Y D A A I D L F T R S C A G Y C V A T F I L G I G D R H N S N 920
CTCAAAGACAAGAACAAGGGGGAAATATATGATGCGGCCATCGATTTGTTTACACGATCATGTGCTGGATATTGTGTTGCCACCTTCATTTTGGGAATTGGAGATCGTCACAATAGTAAT 2760
J K
I M V K D D G Q L F H I D F G H F L D H K K K K F G Y K R E R V P F V L T Q D F 960
ATCATGGTTAAAGATGATGGACAACTGTTTCATATAGATTTTGGACACTTTTTGGATCACAAGAAGAAAAAATTTGGTTATAAACGAGAGCGCGTGCCGTTTGTTTTGACACAAGATTTC 2880
L
L I V I S K G A Q E C T K T R E F E R F Q E M C Y K A Y L A I R Q H A N L F I N 1000
TTAATAGTGATTAGTAAAGGAGCCCAAGAATGCACAAAGACAAGAGAATTTGAGAGGTTTCAGGAGATGTGTTACAAGGCTTATCTAGCTATTCGGCAGCATGCCAATCTCTTCATAAAT 3000
M
L F S M M L G S G M P E L Q S F D D I A Y I R K T L A L D K T E Q E A L E Y F M 1040
CTTTTCTCAATGATGCTTGGCTCTGGAATGCCAGAACTGCAATCTTTTGATGATATTGCATACATTCGAAAGACCCTAGCTTTAGATAAAACTGAGCAAGAGGCTTTGGAGTATTTCATG 3120
N
K Q M N D A H H G G W T T K M D W I F H T I K Q H A L N * 1069
AAACAAATGAATGATGCACACCATGGTGGCTGGACAACAAAAATGGATTGGATCTTCCACACAATTAAGCAGCATGCTTTGAACTGA 3207
(S) EcoNI EcoRI
BstXI
Xba I
PvuII
PstI
HinDIII
AAA n
p2.1
Race Product
a
3634 base pairs

Fig.1

```
AAAAATCTCA ACACATGTGA ATGATCAGAA AATTATCGCC ATAAAAGACA GAATAAGTCA      60
TCAGCGGTTG TTTCATTTCC TATATTTTTT TTTTATTTTT TTATTTTTTA ATAAGGGAAA     120
ATTTAACGTC TAAGGATACA GAAGATTGTT AGCACATTAA AGTAATAAAG GCTTAAGTAG     180
TAAGTGCCTT AGCATGTTAT TGTATTTCAA AGGACATAAT CTAAAATAAT AACAATATCA     240
TTTCTCACAA GTTATTCAAT TTTCTTTTTT TTTTCTAATA ATATCAAGAA TGTATTATTT     300
GTTTGACATA AGTCAACTAA TTTATTTAAT ATGCTGGATT AATCTTGCAG ACATGTAAAT     360
TAACAAGTTT TAGTCAAATA ACGTTGAAGT TTCAATGAAC TCAAATAATT TCTCTTTTTT     420
TTTATATAAC CATATGTCTA ATCTGATTTA TATTTTCCGC AGGATCAACT GAAGTTATGA     480
CATTTGGATT GGATCACTTA TAACCTTGGT CGCCAAATAA TACAAAAATC AGCGTTATAA     540
AACAAAGAAG GTTTTTGTTA AGAAATTAAT CCTCTTTCTT GATAAGAAAG TTGAACCGAA     600
ATTGCAGATA CTGATATATG AAAATAATAC CCACAATTTT GGGAATAGCG CAAGCCTCAA     660
TTTAAACAAT AGGTGAGGAC ACATGATAAT GACCTCAATG ATTGTTAGAA GAAAAGAGCC     720
TCATTACAAA ATCGAAAAAT GAATGGTTGG GTACAAGTTT CCAAAACATG GTAAAGTGGA     780
CTTTGCGTAT GAGACGTAAA TAGAAAAAAA CACTTGTTAT ATGTTTTCTA GAATTATTGT     840
TGTCTCTTTA TGGTTGGATG ATGCAAAATA GTAATTTCGG TTAGTTGCTG TAAAACACCA     900
CGAGACAAAT AGATATGGAT ATTTATTAAA TCAGGAAAAA CGTAACTCTC GGCTACTGGA     960
TGGTTCAGTC ACCCAACGAT TACTGGGGAG AGAAAACAGG GCAAAAGCAA AGCTTAAAGG    1020
AATCCGATTG TCATTCGGCA ATGTGCAGCG AAACTAAAAA CCGGATAATG GACCTGTTAA    1080
TCGAAACATT GAAGATATAT AAAGGAAGAG GAATCCTGGC ATATCATCAA TTGAATAAGT    1140
TGAATTAATT ATTTCAATCT CATTCTCACT TTCTGACTTA TAGTCGCTTT GTTAAATCAT    1200
AGGAATGTCT CCCTTGCCAG TACTGCTAGG GTTTTTCTTT CAAACTATGG AAGCCCATTC    1260
AAGCTGCATA TTACGATTTT GTTTTTCGCT TTTAGAAAGT GGTTTAGATG AGATAATAGA    1320
AAAATTCTTG ATCTCCGACA ACGAGTACTT TTATTTTTTT TGCTAATCAC TTTACTCAAT    1380
ATTAGCTCGA AATCGTAGAA ACGTAGACGG GTGCGGGATA CCGAGTGGTG TAGTTAAGAA    1440
TTTTTATAAA CCACGTGGCC CAAAAATATG AACCCAAAAC GTTTATACAT GAGTATACTT    1500
TAAGAAGGCT ATACCCCTTC GTGTTAGATG TAGTTTTAGC TACCCAACCC GAGTCTATGA    1560
GCTTGACTTC AGATGTAGAA GGCATTAAAT CGTTTTGAAT ATTAATTAAA AAACGATGAA    1620
AATTAAATAT TTAAAAGCAA TCATACGCTG AAAATTTAGT GCTGTGGCTA ATCCTTCAAC    1680
ATGGAAATGC CATAAAAGTG ACTTTGACAA AAAAAAAAGT ATATACAGGT AGTAAACTCA    1740
TCTACTTCAT TGACTTTGTT TACAGCATGT GGAAGGAGGA ATATTTATTG CTAAATCGTA    1800
GTTTAACATT CAATAAGTAA TACTATTGAA ATTCGACAAG ATTGGCCGCA TGGATGAAAA    1860
AGAGGCATTT TGCTTTGGGA GAATTAGTTC AAATTAGAAC TGAAAAAAAA AACTTTACGA    1920
GGCAAAAATG TCGGATTGAG ATCGTAAAAG TTCGCTCGTC GTCTTTTGCT TTGTGATTGT    1980
TTTCATGGAT ACATCTTGCT GGATATTTAA ATTTTAGTAC TATGTATAAG ATATTCTATA    2040
```

Figure 2; page 1 of 2

```
AATGTTTTAT CACCCAAACC TGTTAGCGCC TTCTTAATTC TATTCAATCT GGCTTTTGCT    2100

CTGAGACTAC TTCTTGGACT TTCACTACTT GTTAGTTATA CGGAATTTGT GTAATTAGAA    2160

GTGAAATAAT CCTTTCTATT AGTAATGCAA ACAAAAATC                        2199
```

Figure 2; page 2 of 2

```
CTCGAGCTGA AGAACCAGCG AGGCGGCGAG GCAGCCCCCG CGGCTTGCAG CGGAGGCGAC      60
AGCTCGTCTC CTGCCGTGGA GGTGTCGCCG GTGGTGGGGG GGAGAGACTT GCTCCAAAAA     120
AACGGACGTC TCCAGCTCTC CCCCCTCCCT GTTTTCCGTT AGGAATCCGG CGAGGAAATA     180
CATGCACTCG CTGAGAATCG GCGGCGCCAG GAGGCAGCGC CACAAGGTGT AGCGAGTGAG     240
TGGGGTGGGG CAAGAGGGGA CCCAGGAGTC CCCCAGGCTC CCGGCGCGCC TGCTCCTGCT     300
CTTCAATCCT GCCCACGGGG CGGACGGAGT GACCCCCGCC CCGACCATGG TAGTGTTCAA     360
TGGCCTTCTT AAGATCAAAA TCTGCGAGGC GGTGAGCTTG AAGCCCACAG CCTGGTCGCT     420
GCGCCATGCG GTGGGACCCC GGCCACAGAC GTTCCTTTTG GACCCCTACA TTGCCCTTAA     480
CGTGGACGAC TCGCGCATCG GCCAAACAGC CACCAAGCAA AAGACCAACA GCCCGGCCTG     540
GCACGATGAG TTCGTCACCG ATGTGTGCAA TGGGCGCAAG ATCGAGCTGG CTGTCTTTCA     600
CGACGCTCCT ATCGGCTACG ACGACTTCGT GGCCAACTGC ACCATCCAGT TCGAGGAGCT     660
GCTGCAGAAT GGGAGCCGTC ACTTCGAGGA CTGGATTGAC CTGGAGCCAG AAGGAAAAGT     720
GTACGTGATC ATCGATCTCT CGGGATCATC GGGTGAAGCC CCTAAAGACA ATGAAGAACG     780
AGTGTTCAGG GAGCGTATGC GGCCAAGGAA GCGGCAAGGG GCTGTCAGGC GCAGGGTCCA     840
CCAGGTCAAT GGCCACAAGT TCATGGCCAC CTACTTGCGG CAACCCACCT ACTGCTCCCA     900
CTGCAGAGAT TTCATCTGGG GTGTCATAGG AAAACAGGGA TATCAATGTC AAGTTTGCAC     960
TTGCGTTGTC CACAAGCGAT GTCATGAGCT CATTATTACA AAGTGCGCTG GGCTGAAGAA    1020
ACAGGAAACC CCTGACGAGG TGGGCTCCCA ACGGTTCAGC GTCAACATGC CCCACAAGTT    1080
CGGGATCCAC AACTACAAGG TCCCCACGTT CTGTGACCAC TGTGGGTCCC TGCTCTGGGG    1140
CCTCTTGCGG CAGGGCTTGC AGTGTAAAGT CTGCAAAATG AATGTTCACC GGCGATGTGA    1200
GACCAACGTG GCTCCCAACT GTGGGGTAGA CGCCAGAGGA ATTGCCAAAG TGCTGGCTGA    1260
CCTCGGTGTT ACTCCAGACA AAATCACCAA CAGTGGCCAA AGGAGGAAAA AGCTCGCTGC    1320
TGGTGCTGAG TCCCCACAGC CGGCTTCTGG AAACTCCCCA TCTGAAGACG ACCGATCCAA    1380
GTCAGCGCCC ACCTCCCCTT GTGACCAGGA ACTAAAAGAA CTTGAAAACA ACATCCGGAA    1440
GGCCTTGTCA TTTGACAACC GAGGAGAGGA GCACCGAGCG TCGTCGGCCA CCGATGGCCA    1500
GCTGGCAAGC CCCGGAGAGA ATGGGGAAGT CCGGCCAGGC CAGGCCAAGC GCTTGGGGCT    1560
GGATGAGTTC AACTTCATCA AAGTGTTGGG CAAAGGCAGC TTTGGCAAGG TCATGTTGGC    1620
GGAACTCAAA GGCAAAGATG AAGTCTACGC TGTGAAGGTC TTGAAGAAGG ACGTTATCCT    1680
ACAAGACGAT GATGTGGACT GCACAATGAC AGAGAAGAGG ATTTTGGCTC TGGCTCGGAA    1740
ACACCCTTAT CTAACCCAAC TCTATTGCTG CTTCCAGACC AAGGACCGCC TCTTCTTCGT    1800
CATGGAATAT GTAAATGGTG GAGACCTCAT GTTCCAGATT CAGCGGTCCC GAAAATTTGA    1860
TGAGCCTCGT TCTCGGTTCT ATGCCGCAGA GGTCACATCG GCCCTCATGT TTCTCCACCA    1920
GCATGGAGTG ATCTACAGGG ATTTGAAACT GGACAACATC CTTCTAGATG CAGAAGGCCA    1980
CTGCAAGCTG GCTGACTTTG GGATGTGCAA GGAAGGGATT ATGAATGGTG TGACAACTAC    2040
```

Figure 3; page 1 of 2

```
CACCTTCTGT GGGACTCCTG ACTACATAGC TCCAGAGATC CTACAGGAGT TGGAGTACGG    2100

CCCCTCAGTG GACTGGTGGG CCCTGGGGGT GCTGATGTAC GAGATGATGG CTGGGCAGCC    2160

CCCCTTTGAA GCTGACAACG AGGACGACTT GTTCGAATCC ATCCTTCATG ATGATGTTCT    2220

CTATCCTGTC TGGCTTAGCA AGGAAGCTGT CAGCATCCTG AAAGCTTTCA TGACCAAGAA    2280

CCCGCACAAG CGCCTGGGCT GTGTGGCAGC GCAGAACGGG GAGGACGCCA TCAAGCAACA    2340

TCCATTCTTC AAGGAGATTG ACTGGGTACT GCTGGAGCAG AAGAAAATCA AGCCCCCCTT    2400

CAAGCCGAGA ATTAAAACCA AAAGAGATGT CAATAACTTT GACCAAGACT TTACGCGGGA    2460

AGAGCCAATA CTTACACTTG TGGATGAAGC AATCATTAAG CAGATCAACC AGGAAGAATT    2520

CAAAGGCTTC TCCTACTTTG GTGAAGACCT GATGCCCTGA GAGGCTGCTT CGGATGGAGG    2580

GAGCTCATGC TGCAAGGACG GTGTTGAGAT ACTCCCAAGC TGCAGAGGCT CCGAAGGTCT    2640

CAACTCCTCC TCCTCCTCCC CCTCCCCAGA GCCCCAGTCC CATGTCCACT CTCTTATTTA    2700

TTGCATT                                                              2707
```

Figure 3; page 2 of 2

```
GGCCCCTGTT CTGCAGAAAG GGGGCTCTGA GGCAGAAGGT GGTCCATGAG GTCAAGAGCC      60
ACAAGTTCAC CGCTCGCTTC TTCAAGCAGC CGACCTTCTG CAGCCACTGC ACTGACTTCA     120
TATGGGGGAT TGGAAAACAG GGTCTGCAAT GTCAAGTCTG CAGTTTTGTG GTTCATCGAC     180
GATGCCACGA GTTTGTGACC TTCGAGTGTC CAGGCGCTGG GAAGGGCCCC CAGACGGACG     240
ATCCCCGGAA CAAGCACAAG TTCCGTCTGC ACAGCTACAG CAGCCCCACC TTCTGCGACC     300
ACTGTGGCTC CCTGCTCTAC GGGCTGGTGC ACCAGGGCAT GAAGTGTTCT TGCTGCGAGA     360
TGAACGTGCA CCGGCGCTGT GTGCGCAGCG TGCCCTCTCT GTGCGGCGTG GACCACACGG     420
AGCGCCGGGG CCGCCTGCAG CTGGAGATCC GGGCGCCCAC TTCCGATGAG ATCCACGTTA     480
CGGTTGGCGA GGCCCGGAAC CTCATCCCAA TGGACCCCAA CGGTCTCTCC GATCCCTATG     540
TGAAGCTGAA GCTCATCCCA GACCCTCGGA ATTTGACCAA GCAGAAGACC CGCACGGTGA     600
AAGCTACGCT AAACCCTGTG TGGAACGAGA CCTTTGTGTT CAACCTGAAG CCGGGGGACG     660
TGGAGCGCCG GCTCAGCGTG GAGGTGTGGG ACTGGGACCG GACCTCCCGA AACGACTTCA     720
TGGGCGCCAT GTCCTTCGGC GTCTCGGAGC TGCTCAAGGC GCCGGTGGAC GGCTGGTACA     780
AGTTACTGAA CCAGGAGGAG GGCGAGTATT ACAATGTGCC GGTGGCTGAC GCCGACAACT     840
GCAACCTCCT CCAGAAGTTC GAGGCCTGTA ACTACCCCCT GGAACTATAC GAGAGGGTGC     900
GGACGGGTCC CTCTTCATCT CCCATCCCCT CCCCATCCCC CAGTCCCACC GACTCCAAGC     960
GCTGTTTCTT CGGGGCCAGC CCTGGACGAC TGCACATCTC CGACTTCAGC TTCCTCATGG    1020
TTCTAGGAAA AGGCAGTTTT GGGAAGGTGA TGCTGGCCGA GCGCCGGGGC TCCGATGAGC    1080
TCTACGCCAT CAAGATCCTG AAGAAAGACG TGATCGTCCA GGATGACGAC GTGGACTGCA    1140
CCCTGGTGGA GAAACGCGTG CTGGCTCTGG GGGGCCGAGG CCCGGGAGGC CGGCCGCACT    1200
TCCTCACCCA GCTTCACTCC ACCTTCCAGA CCCCGGATCG CCTGTATTTT GTGATGGAGT    1260
ATGTCACCGG GGGCGACTTG ATGTACCACA TTCAACAGCT GGGCAAGTTT AAGGAACCCC    1320
ACGCAGCGTT CTACGCTGCA GAAATCGCCA TCGGCCTCTT CTTCCTCCAT AACCAGGGCA    1380
TTATCTATCG GGACCTGAAA CTGGACAACG TGATGCTGGA TGCCGAAGGA CACATCAAAA    1440
TCACCGACTT CGGCATGTGT AAGGAGAACG TCTTTCCCGG GAGTACCACT CGCACCTTCT    1500
GCGGGACCCC GGACTACATA GCCCCCGAGA TCATTGCCTA CCAACCCTAT GGGAAGTCTG    1560
TGGATTGGTG GTCCTTTGGG GTTCTGCTCT ACGAGATGTT GGCAGGACAG CCCCCCTTTG    1620
ATGGAGAAGA TGAGGAGGAG CTGTTTCAAG CCATCATGGA ACAAACTGTC ACCTACCCCA    1680
AGTCGCTTTC CCGGGAAGCT GTGGCCATCT GCAAGGGGTT CCTCACCAAG CACCCGGCCA    1740
AGCGCCTGGG CTCAGGCCCC GATGGAGAGC CCACCATCCG CGCTCACGGC TTTTTCCGCT    1800
GGATCGACTG GGACAGGCTG GAACGATTAG AGATCGCGCC TCCGTTCAGA CCCCGCCCGT    1860
GTGGCCGCAG CGGCGAGAAC TTCGACAAGT TCTTCACTCG GGCGGCGCCG GCGCTGACAC    1920
CCCCTGACCG CCTGGTTCTG GCCAGCATCG ACCAGGCTGA GTTCCAGGGC TTCACCTATG    1980
TCAACCCGGA TTTCGTGCAC CCGGATGCCC GCAGCCCCAT CAGCCCAACG CCTGTGCCAG    2040
```

Figure 4; page 1 of 2

```
TCATGTAATC CCACCTGCCG CCACCAGGCG TCCCCACGGC TCCCTCCTCC GCCCCGGCTT    2100

TGGCCCTCGC CTCACCATGC CACCCGCCTT TCCAATTCTA GATATGGCTC CCCAGCGTTC    2160

TGGCCTC                                                              2167
```

Figure 4; page 2 of 2

```
GGCGGCGGCC GCGGGGATCC CGCGAGCGGC CCCTGAACAT CTACCCTTCT TGCCGGGACC      60
CGGGAGGTCC CCACTGGCCT CCGGGCCCGT CCTGATCAGA CTCGTGTCGA CCTCCCCGTC     120
CACGCGCATC CGGGAGAGCC GCGCCACGAG ACGGACCCGG GCCCGCCGGG ACCCCTGGTG     180
TCTGGCCCTG CGTCGAGAGG CTGGTGACTG CCACCCATAA GCTCCAGCTT CAGCCTCGGC     240
TTACTCCCCT CAGGGGCTTG CAGGCTGAGG CCTGCCCTCG GACGCGGCTG ACCAGCCTCT     300
CCCTCTCTTC CACACTTTGG ACTTCTCTTT GGACCTCCTA AAAAGGCTCC ATCATGGCAC     360
CGTTCCTGCG CATCTCCTTC AATTCCTATG AGCTGGGCTC CCTGCAGGCG GAGGACGACG     420
CAAGCCAGCC TTTCTGTGCC GTGAAGATGA AGGAGGCACT CACCACAGAC CGAGGGAAGA     480
CTCTGGTACA GAAGAAGCCC ACAATGTACC CTGAGTGGAA GTCAACATTC GACGCCCACA     540
TCTATGAAGG CCGTGTCATC CAGATCGTGC TGATGCGGGC AGCTGAAGAC CCCATGTCGG     600
AGGTGACCGT GGGCGTGTCA GTGCTGGCTG AGCGCTGCAA GAAGAACAAC GGCAAGGCTG     660
AGTTCTGGCT GGACCTGCAG CCTCAGGCCA AGGTGCTGAT GTGTGTGCAG TATTTCCTGG     720
AGGATGGGGA TTGCAAACAG TCCATGCGTA GTGAGGAGGA GGCCATGTTC CCAACTATGA     780
ACCGCCGTGG AGCCATTAAA CAGGCCAAGA TTCACTACAT CAAGAACCAC GAGTTCATCG     840
CCACCTTCTT TGGGCAGCCC ACCTTCTGTT CTGTGTGCAA AGAGTTTGTC TGGGGCCTCA     900
ACAAGCAAGG CTACAAATGC AGGCAATGCA ACGCTGCCAT CCATAAGAAA TGCATCGACA     960
AGATTATCGG CCGCTGCACT GGCACTGCTA CCAATAGCCG GGACACCATC TTCCAGAAAG    1020
AACGCTTCAA CATCGACATG CCTCACCGAT TCAAGGTCTA TAACTACATG AGCCCCACCT    1080
TCTGTGACCA CTGTGGCACT TTGCTCTGGG GATTGGTGAA ACAGGGATTA AAGTGTGAAG    1140
ACTGCGGCAT GAATGTGCAC CACAAATGCC GGGAGAAGGT GGCCAACCTG TGTGGTATCA    1200
ACCAAAAGCT CTTGGCTGAG GCCTTGAACC AAGTGACCCA GAAAGCTTCC CGGAAGCCAG    1260
AGACACCAGA GACTGTCGGA ATATACCAGG GATTCGAGAA GAAGACAGCT GTCTCTGGGA    1320
ATGACATCCC AGACAACAAC GGGACCTATG GCAAGATCTG GGAGGGGAGC AACCGGTGCC    1380
GCCTTGAGAA CTTCACCTTC CAGAAAGTAC TTGGCAAAGG CAGCTTTGGC AAGGTACTGC    1440
TTGCAGAACT GAAGGGCAAG GAAAGGTACT TTGCAATCAA GTACCTGAAG AAGGACGTGG    1500
TGTTGATCGA CGATGACGTG GAGTGCACCA TGGTGGAGAA GCGGGTGCTG GCGCTCGCCT    1560
GGGAGAATCC CTTCCTCACC CATCTCATCT GTACCTTCCA GACCAAGGAC CACCTCTTCT    1620
TTGTGATGGA GTTCCTCAAT GGGGGCGATC TGATGTTCCA CATTCAGGAC AAAGGCCGCT    1680
TCGAACTCTA CCGGGCTACG TTTTATGCAG CTGAGATCAT CTGCGGACTG CAGTTTCTAC    1740
ATGGCAAAGG CATCATTTAC AGGGACCTCA AGCTAGACAA TGTAATGCTG GACAAGGATG    1800
GCCACATCAA GATTGCTGAC TTCGGGATGT GCAAAGAGAA TATATTTGGG GAGAACCGGG    1860
CCAGCACATT CTGCGGCACT CCTGACTACA TCGCCCCTGA GATCCTGCAG GGCCTGAAGT    1920
ACTCATTTTC CGTGGACTGG TGGTCTTTTG GGGTCCTCCT CTATGAGATG CTCATTGGCC    1980
AGTCCCCCTT CCATGGTGAT GATGAGGACG AGCTCTTTGA GTCCATCCGG GTGGACACAC    2040
```

Figure 5; page 1 of 2

```
CACACTACCC GCGCTGGATC ACCAAGGAGT CCAAGGACAT CATGGAGAAG CTCTTCGAGA    2100
GGGACCCTGC CAAGAGGCTG GGAGTAACAG CAAACATCAG GCTTCACCCC TTTTTCAAGA    2160
CTATCAACTG GAACCTGCTG GAGAAGCGGA AGGTGGAGCC CCCCTTTAAG CCCAAAGTGA    2220
AATCCCCTTC AGACTACAGC AACTTTGACC CAGAGTTCCT GAATGAGAAA CCCCAACTTT    2280
CCTTCAGTGA CAAGAACCTC ATCGACTCTA TGGACCAGAC AGCCTTCAAG GGCTTCTCCT    2340
TTGTGAACCC CAAATATGAG CAATTCCTGG AATAGTGAGC TCCCAGACCT GCTTTTAATG    2400
CCCCGGCAGA GTAGGCCCAT CTGCCCTGGT TTGCATCCTC ACTGCCCATG AAGAAGAGTG    2460
GGTGACTGGT GATTCCTGCT GCTGCCCCCT CTTCCTCGGA GAGTCTGGCT CCTGTTGGCT    2520
GGGCTCACAG TACTTCCTCT GTGAACTGTT TGTGAATTTG CCTTCCTTTT GCCATCGGAG    2580
GGAAACTGTA AATCCTGTGT GTCATTACTT GAATGTAGTT ATTGAAATAT ATATTATATA    2640
TATGCACATA TATATAATAG GCTGTATATA TTGCTCAGTA TAGAAAGCAT GTAGGAGACT    2700
GGTGATGTGT TGACCTTTTT TTAAAAAAAA CCATATGTAT ACGTGTGTAT GTATACATCT    2760
ACACACGTAT ACATATATGT ATGTATGTAT GTATGTATGT ATGTATATAT GACCAAAAGA    2820
AAAGAGAGCA CAAGCTACCT GAACCACAGG ATTGTTTATG TGTGTATAAA TAAACACTGA    2880
ATGGTAAAAA A                                                         2891
```

Figure 5; page 2 of 2

```
TCCGGGTTCC CCAGTGCCAG CCAGCGCGGC CCCCTCGGGG CTCCGGCAGC AGCGCCGGCA      60
TGTCGTCCGG CACGATGAAG TTCAATGGCT ATCTGAGGGT CCGCATCGGA GAGGCTGTAG     120
GGCTGCAGCC CACCCGCTGG TCCCTGCGGC ACTCGCTCTT CAAAAAGGGC CACCAGCTGC     180
TGGACCCCTA CCTGACGGTG AGCGTAGACC AGGTACGCGT GGGCCAGACC AGCACAAAGC     240
AGAAGACCAA CAAACCCACC TACAACGAGG AGTTCTGCGC CAATGTCACC GACGGCGGCC     300
ACCTGGAGCT AGCCGTCTTC CACGAGACGC CCCTGGGTTA TGACCACTTT GTGGCCAACT     360
GCACGCTGCA GTTCCAGGAG CTGTTGCGCA CGGCTGGTAC CTCGGACACC TTCGAGGGCT     420
GGGTGGATCT GGAGCCTGAG GGGAAAGTGT TTGTGGTAAT AACCCTAACA GGGAGTTTCA     480
CTGAAGCCAC TCTCCAGAGA GACCGCATCT TCAAGCATTT TACCAGGAAG CGCCAAAGGG     540
CTATGCGAAG ACGAGTCCAT CAAGTGAACG GACATAAGTT CATGGCCACG TACCTGAGGC     600
AGCCCACCTA CTGCTCTCAT TGCCGAGAGT TCATCTGGGG AGTATTTGGG AAACAGGGTT     660
ATCAATGCCA AGTGTGCACC TGCGTCGTCC ATAAACGCTG CCATCATCTA ATTGTTACAG     720
CCTGCACTTG CCAAAACAAT ATTAACAAAG TGGATGCCAA GATTGCAGAA CAGCGGTTTG     780
GCATCAACAT CCCACACAAG TTCAACGTTC ACAATTACAA GGTGCCCACG TTCTGTGACC     840
ACTGTGGCTC CCTGCTCTGG GGGATAATGC GACAAGGACT TCAGTGTAAA ATATGTAAGA     900
TGAATGTACA TATTCGGTGT CAGGCGAACG TGGCCCCAAA CTGCGGGGTG AATGCCGTGG     960
AGCTTGCCAA GACCCTGGCA GGGATGGGTC TCCAACCCGG AAATATTTCT CCAACCTCGA    1020
AACTCATTTC CAGGTCGACA CTAAGACGGC AGGGAAAGGA GGGCTCCAAA GAAGGAAATG    1080
GGATCGGTGT TAACTCTTCC AGCAGATTCG GCATCGACAA CTTTGAGTTC ATCCGGGTGT    1140
TGGGGAAGGG GAGCTTCGGG AAGGTGATGC TTGCCAGGAT AAAGGAGACA GGAGAACTGT    1200
ACGCCGTGAA GGTGCTGAAG AAGGACGTGA TTCTGCAGGA TGATGATGTA GAGTGCACCA    1260
TGACTGAGAA GAGGATCCTG TCCTTGGCTC GCAACCACCC CTTCCTCACC CAGCTCTTCT    1320
GCTGCTTTCA GACTCCAGAC CGTCTGTTCT TTGTCATGGA GTTTGTGAAC GGAGGCGACC    1380
TGATGTTCCA CATCCAAAAG TCCCGTCGTT TCGATGAAGC CCGTGCTCGT TTCTACGCCG    1440
CGGAGATCAT TTCTGCACTC ATGTTCCTAC ATGAGAAAGG TATCATCTAT AGAGACTTGA    1500
AACTGGACAA TGTGCTATTG GACCACGAAG GTCACTGTAA ACTGGCCGAT TTTGGAATGT    1560
GCAAGGAGGG GATTTGTAAT GGGGTCACCA CAGCCACCTT CTGCGGTACA CCTGACTACA    1620
TTGCCCCAGA GATCCTTCAG GAGATGCTGT ATGGACCTGC AGTAGACTGG TGGGCCATGG    1680
GCGTGTTGCT TTATGAGATG CTGTGCGGAC ATGCGCCCTT TGAGGCTGAA AATGAAGATG    1740
ACCTTTTTGA GGCCATACTG AATGATGAAG TCGTCTACCC CACCTGGCTC CATGAAGATG    1800
CCAGAGGGAT CCTCAAGTCT TTCATGACCA AGAACCCCAC CATGCGCTTG GGCAGCCTGA    1860
CTCAGGGAGG AGAGCATGAG ATCCTGAGAC ACCCTTTCTT TAAGGAAATC GACTGGGCCC    1920
AGTTGAACCA TCGCCAGTTA GAGCCGCCTT TCCGACCTAG AATCAAATCC CGAGAAGATG    1980
TCAGCAATTT TGACCCAGAC TTTATAAAAG AAGAGCCCGT CTTAACTCCG ATTGATGAGG    2040
```

Figure 6; page 1 of 2

```
GACATCTTCC TATGATTAAC CAGGATGAGT TTAGAAACTT TTCCTATGTG TCACCGGAAT    2100

TGCAACTGTA GCCTTATGGG GAGTCAGAAC CAAAGGGGAA GGTGGATTTC TCCAGGAATT    2160

TCTTATGTGG GAATTC                                                   2176
```

Figure 6; page 2 of 2

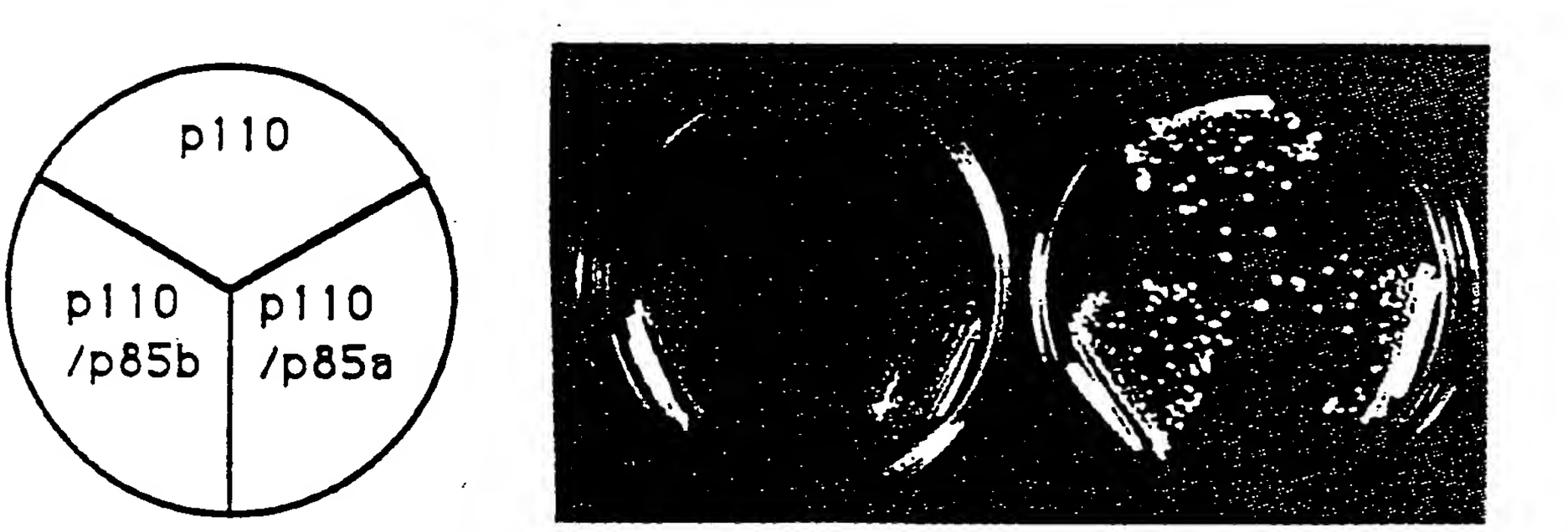
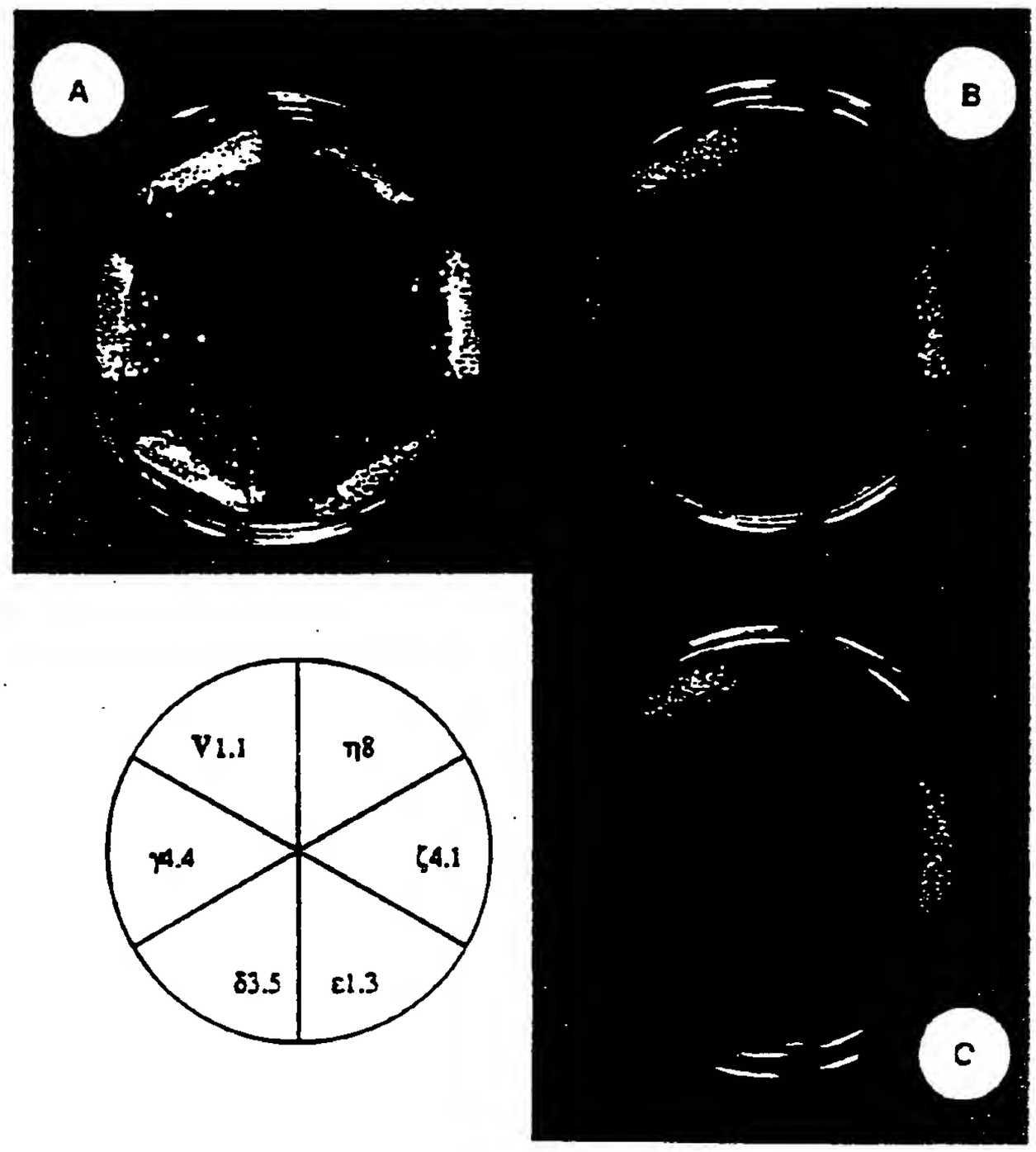


Fig.7



A = 10nM Thiamine
B = n11
C = 10ng/ml TPA

Fig. 8

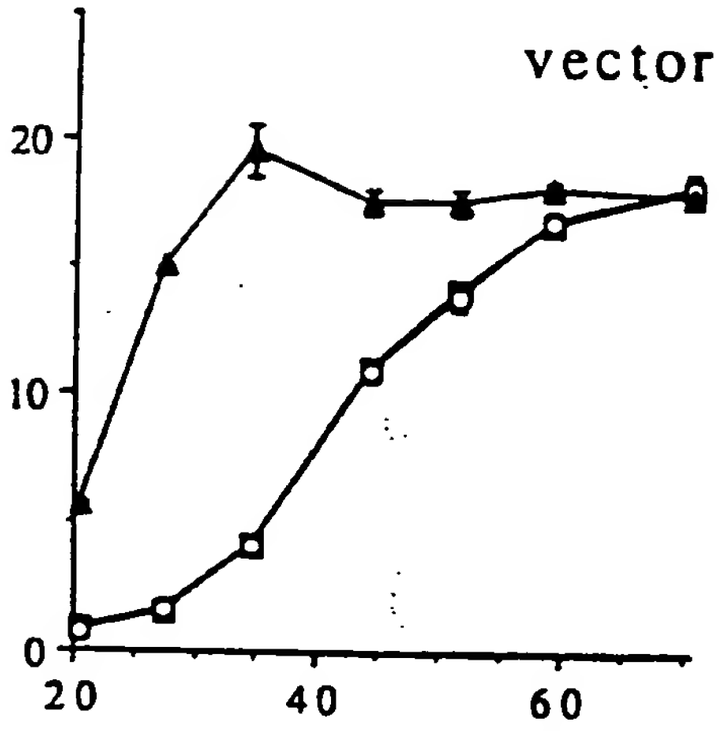


Fig.9A

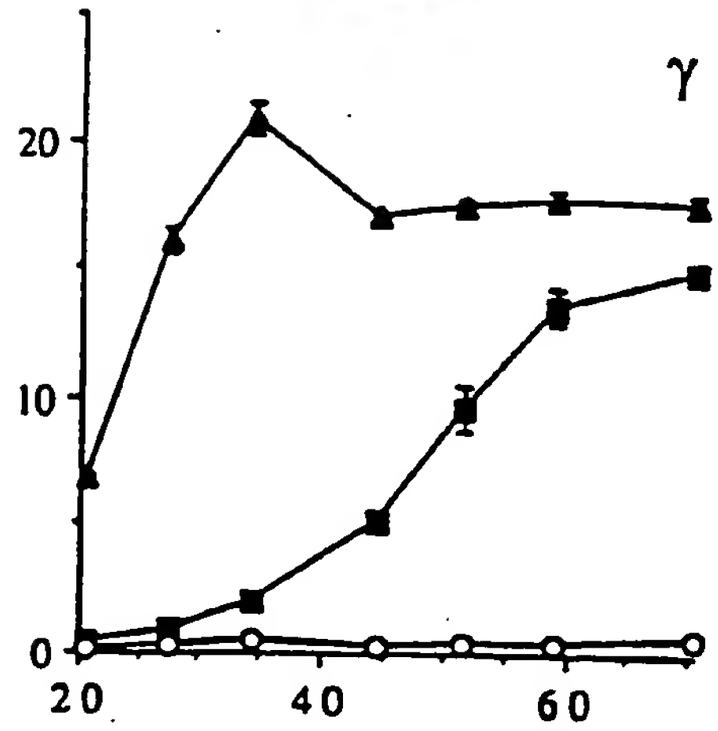


Fig.9B

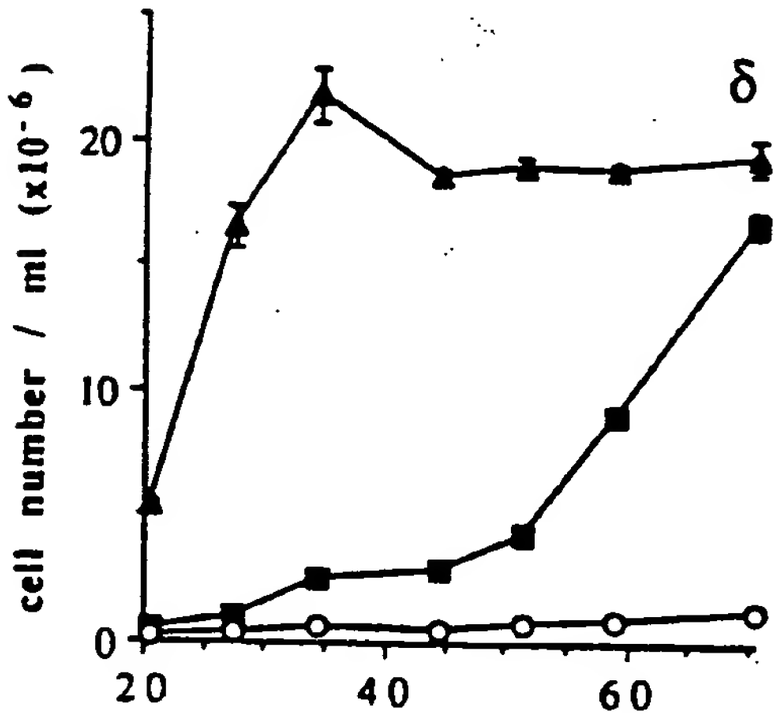


Fig.9C

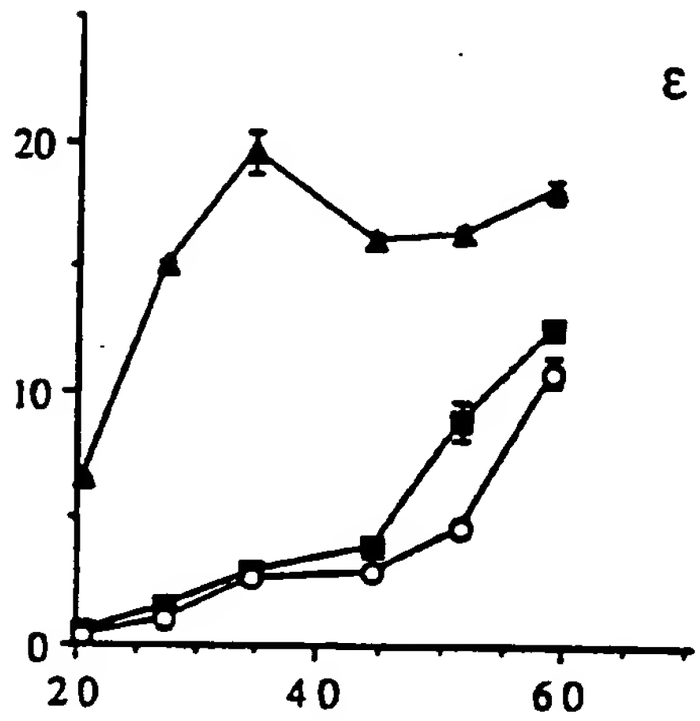
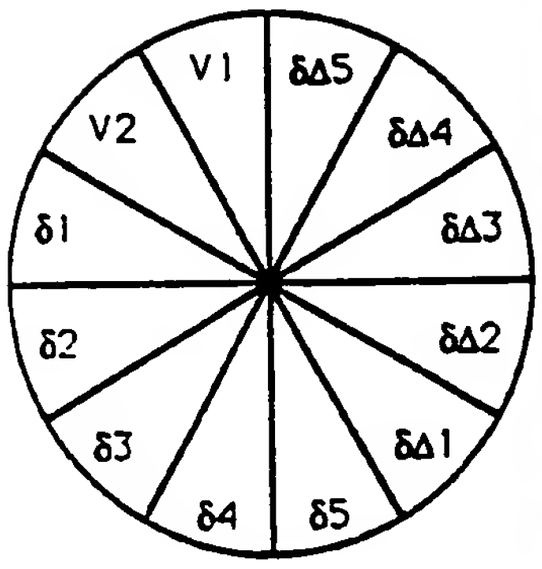


Fig.9D

Fig.9E

Fig.9F


V1
δΔ5
V2
δΔ4
δ1
δΔ3
δ2
δΔ2
δ3
δΔ1
δ4
δ5

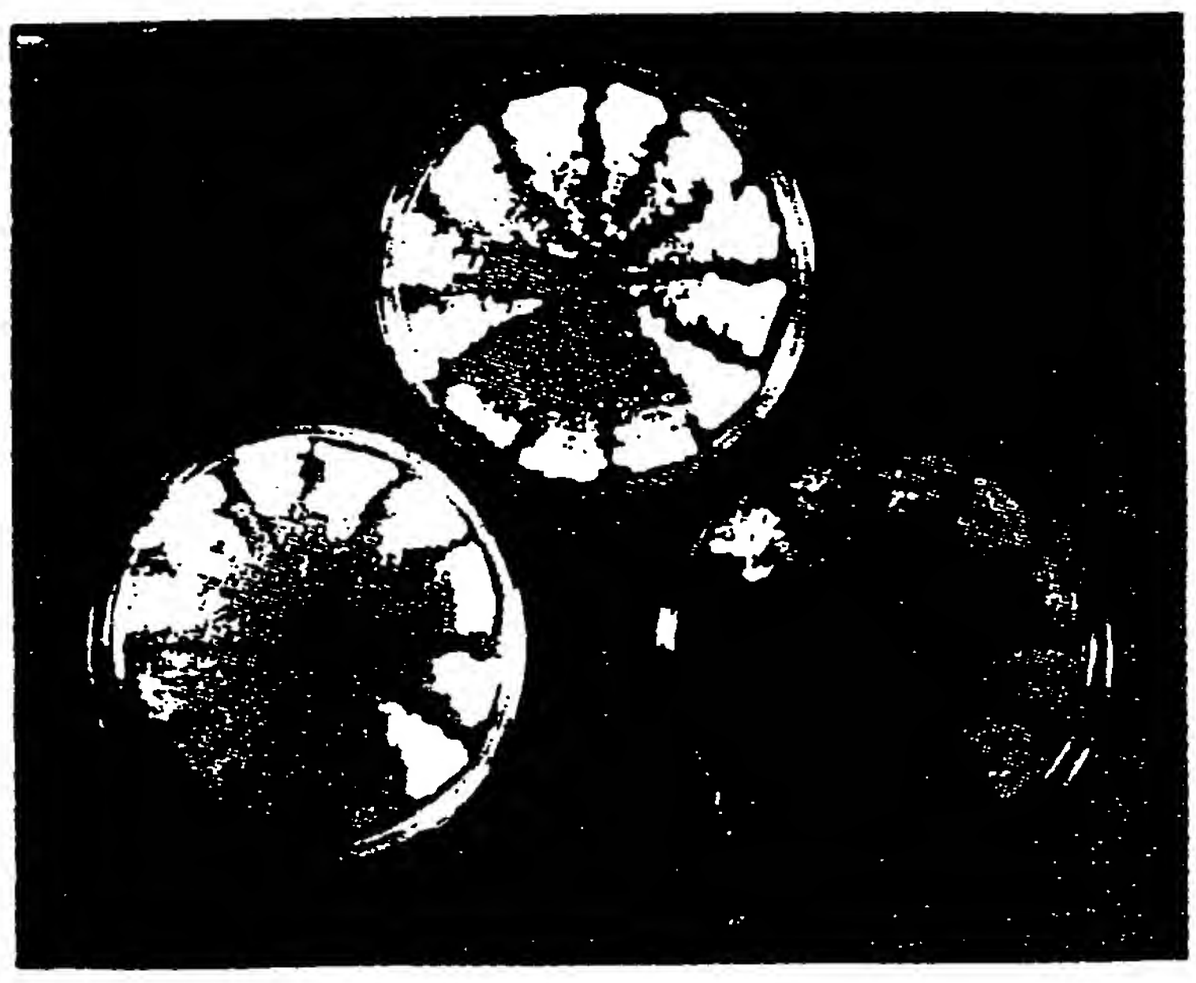

+ Thiamine
- Thiamine
- Thiamine + TPA

Fig. 10

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 93/01651

A. CLASSIFICATION OF SUBJECT MATT[redacted]
IPC 5 C12N15/54 C12N[redacted] C12Q1/48 C12N1/19 [redacted]Q1/02,
(C12N1/19,C12R1:645)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 5 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP,A,0 358 325 (TAKEDA CHEMICAL INDUSTRIES LTD.) 14 March 1990 see page 3, line 14 - line 34	1,2,8, 12-15
Y	see page 4, line 28 - line 36 see example 4	1-6,8,11

Y	JOURNAL OF BIOLOGICAL CHEMISTRY vol. 265, no. 19 , 5 July 1990 , BALTIMORE, MD US pages 10857 - 10864 MAUNDRELL, K. 'nmt1 of fission yeast' cited in the application see from page 10860, right column, last paragraph to page 10864 see figure 8	1-6,8

	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 November 1993	30 -11- 1993
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	**Authorized officer** ANDRES, S

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int[redacted]nal Application No
PCT/GB[redacted]/01651

C.(Continuation) DOCUMENTS CONSIDER[redacted] BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JOURNAL OF BIOLOGICAL CHEMISTRY vol. 265, no. 36 , 25 December 1990 , BALTIMORE, MD US pages 22434 - 22440 OSADA, S.-I. ET AL. 'A phorbol ester receptor/protein kinase, nPKCeta, a new member of the protein kinase C family predominantly expressed in lung and skin' see the whole document	1,8,11 11
Y	---	
X	WO,A,88 01303 (GENETICS INSTITUTE, INC.) 25 February 1988 see page 4, line 12 - page 5, line 33 see page 12, line 32 - page 13, line 7 see example VIII	1,2,8, 12-15

X	WO,A,89 07654 (PROGENICS PHARMACEUTICALS, INC.) 24 August 1989 see page 14 - page 16, line 23 see page 19 - page 22, line 3	1,8, 12-15

P,X	CELL vol. 70 , 7 August 1992 , CAMBRIDGE, NA US pages 419 - 429 HILES, I. ET AL. 'Phosphatidylinositol 3-kinase: structure and expression of the 110 kd catalytic subunit' cited in the application see the whole document	1,7,9,10

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/G[redacted]/01651

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0358325	14-03-90	JP-A-	2138983	28-05-90
WO-A-8801303	25-02-88	AU-A-	7851387	08-03-88
		EP-A-	0317574	31-05-89
		JP-T-	2500243	01-02-90
WO-A-8907654	24-08-89	US-A-	4980281	25-12-90
		AU-B-	612948	18-07-91
		AU-A-	3184089	06-09-89
		EP-A-	0403506	27-12-90
		JP-T-	3503598	15-08-91